Siglo XXI: bibliotecas digitales con Greenstone

Colectivo de Autores



QA402.3 S577 2013

Fresno Chávez, Caridad

Siglo XXI: bibliotecas digitales con Greenstone / Caridad Fresno Chávez ... [et al.]. -- La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0. -- 281 pág.

1. Vázquez Velázquez, Elvia
2. Computadoras - Aspectos sociales - Interacción hombre-computadora
3. DCC 003.5 - CDU 7

Depósito Legal: 9789591614582

Autores: Caridad Fresno Chávez; Elvia Vázquez Velázquez; Ricardo Fernández Rodríguez; Raul G. Torricella Morales y Dailyn Rodríguez Martínez.

Editora: Dra. C. Caridad Fresno Chávez, (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM)

Digitalización: Dr. C. Raúl G. Torricella Morales (torri@reduniv.edu.cu). Director de la Editorial Universitaria.



Portal de la Editorial Universitaria: http://eduniv.mes.edu.cu

Dirección postal: Calle 23 esq. F. No. 564. El Vedado, La Habana, CP 10400, Cuba.

Colectivo de Autores



Dra. C. **Caridad Fresno Chávez** (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH). **Capítulos: 1, 2, 3, 4, 5, 6, 7, 11, 12, 13, 18.**



Lic. **Elvia Vázquez Velázquez** (elvia@cecam.sld.cu) Licenciada en Letras. Centro de Cibernética Aplicada a la Medicina (CECAM) I.S.C.M.H. Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. **Capítulo: 8, 10, 15, 16, 17.**



Ms C **Ricardo Fernández Rodríguez** (ricardo@cecam.sld.cu) Profesor Auxiliar del I.S.C.M.H. Máster en Informática en Salud. Centro de Cibernética Aplicada a la Medicina (CECAM) I.S.C.M.H. Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba. CP: 11600, Telf.: (537) 271-1354. Capítulos: **4, 5, 13.**



Dr. C. **Raul G. Torricella Morales** (torri@reduniv.edu.cu). Prof. Titular adjunto, Facultad de Farmacia y Alimentos de la Universidad de La Habana, Director de la Editorial Universitaria (Cuba). Dirección: Calle 23 # 525 Esq. F., Municipio Plaza, Ciudad Habana, Cuba, CP: 1400, Telf.: (537) 830-3674. Ministerio de Educación Superior de la República de Cuba. **Capítulo 9.**

Tec. **Dailyn Rodríguez Martínez**. Técnico Medio en Informática. Estudiante de 5to año de la carrera de Bibliotecología y Ciencia de la Información. **Capítulo 18.**

Prefacio

La intención de este libro es constituirse en un texto básico apropiado para aquellos interesados en la utilización del Software Greenstone para la creación de Bibliotecas Digitales.

Es por ello que resulta característico en el mismo abordar en cada capítulo con un lenguaje muy diáfano toda la información técnica necesaria para proporcionar la asimilación del conocimiento suficiente para manipular esta herramienta.

Por tanto hemos adoptado un enfoque y distribución de contenidos instructivos de forma gradual, en el nivel de complejidad y de asequibilidad, utilizando muchas ilustraciones que avalen los conceptos e ideas que en el se incluyen.

También contará el lector interesado con la bibliografía necesaria para lograr una consolidación y ampliación de los conocimientos brindados.

En este libro han colaborado un grupo de compañeros que han querido volcar toda su experiencia en la utilización de este software.

A todos ellos, muchas gracias.

Los autores.

Tabla de Contenidos

Prólogo de la primera edición en español

Greenstone es un sistema de código abierto (*Open Source*) creado por la Universidad de Waikato, con el propósito de facilitar a los interesados, la creación de bibliotecas digitales o repositorios institucionales.

Los autores de este libro han intentado volcar en el, toda la experiencia acumulada en el desarrollo de su propio proyecto de creación de una biblioteca digital, y la experiencia alcanzada como docentes, en el dictado de cursos básicos y avanzados en esta materia, con el objetivo de promover la utilización del mismo en Cuba y en Latinoamérica como parte del Proyecto Regional "Grupo de Usuarios Greenstone de Latino-américa" (http:www.greenstone-la.org).

La gran actividad demostrada por el Centro de Cibernética Aplicada a la Medicina (CECAM) en pos de la difusión de Greenstone no hace más que manifestar el profundo deseo que tienen por hacer del acceso a la información uno de los pilares de la sociedad.

Estimamos que esta obra en español, puede constituirse en un valioso aporte docente y práctico.

Lic. Diego Spano
Creador del Grupo de Promoción y soporte
de Greenstone para Latino-américa
Administrador de la lista de usuarios en español

SECCIÓN I: LAS FORMAS BÁSICAS DE TRABAJO CON GREENSTONE

"Pero uno no sabe lo que sabe -ni siquiera lo que desea saber- hasta que lo desafían y se ve obligado a hacerle frente".[1]

Thornton Wilder, 1897-1975.

[1] Tomado de: Saavedra Saldívarel, Justino Ricardo (2010). Etiquetado social de objetos digitales en bibliotecas universitarias: prácticas recomendables. Tesis Maestría (Maestría en Bibliotecología y Estudios de la Información)-UNAM, Facultad de Filosofía y Letras, 2010.

CAPÍTULO 1

1- La era de la información y el conocimiento

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH)

Palabras clave: Alfabetización Digital, Alfabetización en Información, Desarrollo Sostenible, Economía Digital, Estrategias de Inclusión Digital. , Sociedad de la Información.

1.1- Introducción

A consecuencia de la nueva distribución geopolítica que provocó la II Guerra Mundial, el avance del proceso de capitalización en la economía, dió lugar al surgimiento de empresas transnacionales y al inicio de una globalización económica. Este fenómeno, unido a las crisis económicas de los años 70 favoreció el surgimiento, aplicación y desarrollo de nuevas tecnologías de información y comunicaciones (1)

Presenciamos entonces la transición del capitalismo mercantil y esclavista, al capitalismo industrial denominado también segundo capitalismo o capitalismo industrial, y constatamos, de acuerdo a algunos especialistas en el tema, que nos encontramos en fase de transito hacia lo que se ha dado en llamar el Tercer Capitalismo o Capitalismo Cognitivo (2) el cual se caracteriza por colocar en primer plano la redefinición de los derechos de propiedad, en una sociedad digital.

En oposición a esta tendencia se encuentra la denominada Sociedad de la Información y el Conocimiento.

Cada vez más se oye hablar de una nueva sociedad en la que el conocimiento ocupe un lugar muy especial; se propone que la era de la información en que vivimos se transforme en la Sociedad o Economía del Conocimiento.

Con el desarrollo logrado en tecnología de información, comunicaciones y en computación, se están transformando las relaciones entre los distintos países del mundo, y por supuesto, también la economía y las formas en que se llevan a cabo las operaciones económicas, dando lugar a una nueva economía. Se espera que esta evolución se mantenga y se incorporen más naciones de forma plena a esta revolución tecnológica y a los beneficios que esto puede representar.

1.2- Contexto histórico-social en que se presenta la era de la Información y el Conocimiento

La Sociedad de la Información, más que un proyecto definido, es una aspiración: la del nuevo entorno humano, en donde los conocimientos, su creación y propagación son el elemento definitorio de las relaciones entre los individuos y entre las naciones. Ha sido planteado por especialistas en el tema que el rasgo mas singular de esta nueva sociedad es su carácter global y local (3).

El elemento global garantizaría los nexos entre naciones y permitiría la colaboración y participación en un marco de pleno goce de los nuevos adelantos tecnológicos que actualmente dividen a nuestros países en pobres y ricos, en lo que se ha dado en llamar brecha digital (4) y el carácter local, que identificaría y conservaría lo realmente autóctono de cada pueblo.

La diferencia de esta sociedad informacional emergente es que entramos ahora en una forma específica de organización social en la que la generación, proceso y transmisión de información se convierten en las principales fuentes de productividad y poder (5).

Sociedad de la Información/Conocimiento que, de acuerdo a los especialistas en el tema, debe ser denominada Sociedad Informacional, para distinguirla de otras sociedades previas en las que la información también ha jugado un papel importante, enfocándonos en los aspectos relacionados con la creación, transferencia y reutilización del conocimiento o ciclo del conocimiento, destacando también las peculiaridades de la cadena de valor del conocimiento que rige el desarrollo de esta nueva economía del conocimiento donde juega un papel primordial el hombre y el desarrollo de un capital intelectual y humano que garantice el éxito de esta economía informacional/economía del conocimiento/ economía digital, como indistintamente suele ser denominada (3)..

En la Sociedad de la Información lo esencial es la explotación inteligente de la información, su conversión en conocimiento. La generación de mecanismos de aplicación de ese conocimiento en el desarrollo de nuevos productos y procesos, incluso de nuevas formas de organización social es, pues, el recurso fundamental, el que explica las diferencias entre la empresa y las sociedades (6).

Los antecedentes de la Sociedad de la Información/Conocimiento son la - Revolución tecnológica y de las Nuevas Tecnologías de Información y Comunicaciones y el fenómeno de la Globalización (1).

El proceso de globalización se puede definir como el ensanchamiento geográfico de los nexos internos de productos, mercados, empresas y factores de producción con un componente mayor del cual fueron derivados, generados o disponibles en otros países y regiones (1).

Representa un proceso de interconexión financiera, económica, política y cultural que se relaciona con la progresiva expansión internacional de las actividades empresariales gestadas a partir de la II Guerra Mundial que cobra fuerza desde 1970 con la crisis económica que promueve la incorporación de la Nuevas Tecnologías Comunicacionales, estiman los especialistas (1).

El fenómeno de integración de las distintas economías no es un fenómeno nuevo, sino que es el resultado de un largo proceso histórico y de la propia lógica de la acumulación del capital. Los especialistas en el tema aseguran que el proceso histórico de despliegue del capitalismo nacido de la revolución industrial tenía dos opciones, o adquiría el rasgo de la globalidad o desaparecía (2).

1.3- Características de la Sociedad de la Información y el Conocimiento

Entre las características de esta nueva sociedad emergente, se destaca la necesaria participación de todos y cada uno de los actores de la misma, lo que posibilitará elevar la calidad de vida de todos los ciudadanos y explotar al máximo la creatividad del hombre.

> *"El término Sociedad de la Información se refiere a una forma de desarrollo económico y social en el que la adquisición, almacenamiento, procesamiento, evaluación, transmisión distribución y diseminación de la información con vistas a la creación de conocimiento y a la satisfacción de las necesidades de las personas y de los organismos, juegan un papel central en la actividad económica, en la creación de*

riqueza y en la definición de la calidad de vida y las prácticas culturales de los ciudadanos". (7).

Ha sido planteado por especialistas en el tema (3, 7, 8) que el rasgo mas singular de esta nueva sociedad es su carácter global y local.

El Informe Bangemann delineó la estrategia que debía seguir la Comunidad Europea para lograr el desarrollo de una Sociedad de la Información en una Economía Digital (9).

En los propósitos de esta estrategia se incluían diez aplicaciones básicas; Tele-trabajo, Educación a Distancia, Servicios Telemáticos para las Pequeñas y Medianas Empresas (PYMES), Gestión del Tráfico por carretera, Control del Tráfico Aéreo, Redes de Asistencia Sanitaria, Licitación Electrónica, Redes de Administraciones Públicas, Autopistas Urbanas de la Información, Redes de Universidades y Centros de Investigación (9).

Posteriormente la iniciativa eEurope: Una Sociedad de la Información para Todos, lanzada en diciembre de 1999 y perfeccionada en el 2002 (10) y a continuación el Plan de Acción eEurope 2005 (11), dieron continuidad al Informe Bangemann y tuvieron expresiones concretas en las recientes Cumbres de la Sociedad de la Información celebradas en 2003 y 2005 (12).

1.4- Desafíos que enfrenta la construcción de la Sociedad de la Información y el Conocimiento

Existen retos hasta ahora infranqueables, que exigen aún del esfuerzo y dedicación de la comunidad mundial de usuarios de esta nueva sociedad que se vislumbra

En la construcción de una Sociedad de la Información / Conocimiento resulta esencial asegurar la creación del conocimiento, compartir y transferir este conocimiento, esto implica vencer una serie de retos entre los que pueden mencionarse:

- Los aún insuficientes mecanismos de búsqueda, y evaluación de la calidad de la información de que disponemos en una sociedad en la que la información se convierte en un recurso estratégico (13) acompañado de un,
- Fenómeno de exceso de información que caracteriza actualmente a nuestra sociedad y que satura nuestros sistemas informativos (5)
- El desarrollo de un diseño adecuado de estrategia de alfabetización digital de todos los ciudadanos, que permita trazar el camino hacia el objetivo de una Sociedad de la Información para todos (14).
- Complementar la alfabetización digital con la creación de un nuevo modelo educativo que permita incorporar las ventajas de las nuevas tecnologías en información y comunicaciones (15) y desarrollar al máximo la creatividad de los estudiantes (16),
- Implementar estrategias de elevación de la calidad de la educación superior (17), promoviendo niveles de excelencia en la formación de los profesores que llevarán a cabo este proceso (18) y de ese modo situarse al nivel del desarrollo tecnológico que exige una Economía Digital, donde resulta imperante un modelo especial de creación de capacidades y modelos curriculares flexibles (19), que ajusten con el vertiginoso desarrollo económico-social que caracteriza esta economía (20).
- La desvinculación del acceso a la Educación Superior de las leyes de oferta y demanda que rigen este proceso con los actuales modelos económicos (21), y
- La remodelación de los procesos sustantivos de la Universidad con la implementación de

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

procesos de internamiento eficientes, con apropiados mecanismos de gestión financiera y económica y por supuesto la vinculación de la Universidad a las realidades de las comunidades en que se insertan (22),

- La creación de una cultura de la innovación y creatividad que asegure el desarrollo de un capital humano en proceso constante de superación (23), y
- Vigorizar el proceso de creación de organizaciones en aprendizaje, que se considera, serán las únicas capaces de asimilar los vertiginosos cambios que se avecinan (24).

Sumemos a estos aspectos, los dilemas que presentan, además, los denominados países del segundo y tercer mundo, y la lista se hace prácticamente interminable, pero a pesar de ello, queremos destacar dos de los problemas más álgidos que constituyen la espina dorsal del actual desarrollo GEO-político económico del mundo:

- La necesidad de mantener la identidad de cada uno de nuestros países y garantizar un desarrollo sostenible en un mundo globalizado y regido por los intereses de empresas transaccionales (25),
- La dificultad en reducir la brecha digital dictada por las diferencias sociales económicas y políticas que sufren las naciones de menos desarrollo (26) y convertirla en inclusión digital (20).

Los nuevos desafíos planteados por la brecha digital y la oportunidad de convertirla en inclusión digital tienen que ver con la posibilidad de utilizar las Nuevas Tecnologías Comunicacionales como instrumento de desarrollo humano, la lucha contra la pobreza, la consolidación democrática, el incremento de la multiculturalidad, la promoción de una cultura de paz, el fortalecimiento y defensa de la libertad de expresión, la integración y el progreso sostenible. La reducción de la brecha digital debe convertirse en una prioridad mundial. Sin una acción determinada, el crecimiento irregular de la economía del conocimiento interconectado que se avizora puede incrementar la desigualdad y sus consecuencias sociales.

1.5- Consideraciones finales

Resulta indispensable considerar como la mejor estrategia para el desarrollo de una Sociedad Mundial de la Información / Conocimiento, la instauración de políticas adecuadas de inclusión digital, incorporando capacidades para el dominio de las nuevas tecnologías de información y comunicaciones o como también se le conoce: alfabetización digital y el desarrollo de Redes de Conocimientos mediante la construcción de bibliotecas digitales puede ser una alternativa viable para nuestros pueblos que contribuya al desarrollo de una plena economía digital, de modo que pueda mitigarse la realidad que nos golpea e impone el desarrollo de una Sociedad de la Información/Conocimiento al menos a dos velocidades como gusta a muchos especialistas en el tema denominar.

1.6- Referencias Bibliográficas

(1) Quéau P. La revolución de la información en la búsqueda de un bien común. ACI [Seriada en Internet] vol.9 no.01, 1997. Disponible en <http:bus.sld.cu/revistas/aci/vol9_s_oi/sci15100.htm>

(2) Boutang Y. Riqueza, propiedad, libertad y renta en el capitalismo cognitivo en: Capitalismo cognitivo, propiedad intelectual y creación colectiva. [Libro en Internet]. Edición: Traficantes de Sueños Madrid; Abril de 2004. Disponible en: <http://traficantes.net7>.

(3) Castells M. The Information Age: Economy, Society, and Culture. Vol.3 End of Millenium. [Book in Internet] Blackwell Publisher (Oxford and Malden,Ma)1998. Available from: <http://sociology.berkeley.edu/faculty/castells/trilogy_reviews.html>.

(4) Gaynor M. Brecha digital- ¿un problema?. La Brecha Digital. Artículos. [Sitio Web en Internet] Disponible en:<http://www.labrechadigital.org/articulos.html>.

(5) Cornella A. En la sociedad del conocimiento, la riqueza está en las ideas [Libro en Internet] en Infonomia.com: La gestión de la información en las organizaciones. Ediciones Deusto. Madrid, 2002. Disponible en: < http://www.infonomia.com/libro/index.asp>.

(6) Chapeta E. El impacto de la TI en la economía .IW-Global. News. [Seriada en Internet] Mayo 2002. Disponible en: <http://www.iw-global.com/downloads/Press/Newsletter/pag1.pdf>.

(7) García M. Sociedad de la Información. Definiciones. Revista Iberoamericana de Ciencia, Tecnología, Sociedad e Innovación. [Seriada en Internet] No.1/Septiembre-Diciembre 2001.Monográfico. Sociedad de la Información. Debate. Servicios y Aplicaciones: Disponible en: <http://www.campus-oei.org/revistactsi/numero1/index.html>.

(8) Kyushu Okinawa. Carta de Okinawa sobre la Sociedad de la Información Global. [Sitio Web en Internet]. Summit G8. Julio de 2000. Disponible en: <http://members.fortunecity.com/adbaorg/g8.htm>.

(9) Bangemann Report. Europe and the global information society. Recommendations to the European Council [Website in Internet]. Brussels. Luxemburg : European Commission, 25-26 June 1994. Available from:<http://www2.echo.lu/eudocs/en/bangemann.html >

(10) eEurope2002. Informe de evaluación comparativa de la acción eEurope. Comunicación de la Comisión al Consejo al Parlamento Europeo, al Comité Económico y Social y al Comité de las Regiones. [Sitio Web en Internet]. COM (2002) 62 final. Bruselas, 5.2.2002. . Disponible en: <http://europa.eu.int/information_society/eeurope/news_library/documents/index_en.htm>.

(11) eEurope2005. eEurope 2005: Una sociedad de la información para todos. Comunicación de la Comisión al Consejo al Parlamento Europeo, al Comité Económico y Social y al Comité de las Regiones. [Sitio Web en Internet]. COM (2002) 263 final. Bruselas, 28.5.2002 Disponible en: <http://europa.eu.int/information_society/eeurope/news_library/documents/index_en.htm>.

(12) Cumbre Mundial sobre la Sociedad de la Información. CMSI. Ginebra 2003-Túnez 2005. Construir la Sociedad de la Información: un desafío global para el nuevo milenio. Declaración de Principios. [Sitio Web en Internet].Documento WSIS-03/GENEVA/4-S. 12 de mayo de 2004. Disponible en:< http://www.itu.int/dms_pub/itu-s/md/03/wsis/doc/S03-WSIS-DOC-0004!!MSW-S.doc>.

(13) Fresno C. La importancia de la calidad de la información en Internet para el desarrollo de una Sociedad Informacional. Revista Cubana de Informática Médica [Seriada en Internet]. No.4

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

Año 5.Disponible en: <http://www.cecam.sld.cu/pages/rcim/revista_9/articulos_htm/calidadinform.htm>.

(14) Bruce S. Las siete caras de la alfabetización en información en la Enseñanza Superior. Anales de Documentación [Seriada en Internet]. 2003, No.6, p.289-294. Disponible en: <http://www.um.es/fccd/anales/ad06/ad0600.html>. Traducción del trabajo original: Bruce S. Seven faces of information literacy in Higher Education [on line].1997. Available from: <http://sky.fit.qut.edu.au/~bruce/inflit/faces/faces1.htm>.

(15) Marqués P. Nueva cultura, nuevas competencias para los ciudadanos. La alfabetización digital. Roles de los estudiantes hoy. [Sitio Web en Internet] Departamento de Pedagogía Aplicada, Facultad de Educación, Universidad Autónoma de Barcelona (UAB) 6/06/05. Disponible en: <http://dewey.uab.es/pmarques/competen.htm>.

(16) Picardo O. Pedagogía informacional: enseñar a aprender en la Sociedad del Conocimiento. CTSI [Seriada en Internet] mayo-agosto 2002, No.3. Disponible en: <http://www.campus-oei.org/revistactsi/numero3/art04.htm>

(17) UNESCO. Declaración Mundial sobre la Educación Superior en el siglo XXI: Visión y Acción. Conferencia Mundial sobre la Educación Superior [Sitio Web en Internet] Paris: 5-9 de octubre 1998. Disponible en: <http://www.unesco.org/education/educprog/wche/declaracion_spa.htm>.

(18) Valdivieso P. El docente en la nueva era de la información. [Sitio Web en Internet] Universidad Técnica Particular de Loja. 19-May-2005. Disponible en: <http://www.utplonline.edu.ec/internas/articulos/docente_virtual.pdf>.

(19) Gómez, V. M., Celis, J. G. Factores de innovación curricular y académica en la Educación [Sitio Web en Internet] Universidad Nacional de Colombia. Sede Medellín. Marzo 2004. Disponible en: <http://www.unalmed.edu.co/~diracad/reforma_academica/ANEXOSLINEAMIENTOS.doc>.

(20) Fresno C. La creación de la Sociedad de la Información/conocimiento. ¿y la brecha digital? Fundación Latinoamericana de la Calidad [Sitio Web en Internet] 14-mayo-2003. Disponible en: <http://www.calidad.org/s/biblio.php3>.

(21) Piñera J. Una salida para América Latina. Diario El Mercurio. [Sitio Web en Internet] 8 de junio de 2003. Disponible en: <http://www.cisle.org.mx/Publicaciones/pub/america.htm>.

(22) Sebastián J. Oportunidades e iniciativas para la cooperación iberoamericana de Educación Superior. Revista Iberoamericana de Educación [Seriada en Internet] Monográfico Enseñanza de la Tecnología, No. 28: enero-abril 2002. Disponible en: <http://www.campus-oei.org/revista/rie28t.htm>.

(23) Amidon D. The Emerging 6th Generation. [Website in Internet] Entovation Internacional-Delivering Knowledge Innovation Strategies for the Millennium. 27Jan 2000. Available from: <http://www.entovation.com/assessment/best.htm>.

(24) SengeP. The Fifth Discipline: The Art and Practice of the Learning Organization. [Book in Internet], 1990. 1st edition 1994 Currency Doubleday- Bantam Doubleday Dell Publishing Group Inc. Review of the Fifth Discipline, J. P. Fullerton, 1996: Available from: <http://www.rtis.com/nat/use/jfullerton/review/review.htm >.

(25) ONU. Objetivos de la ONU de Desarrollo del Milenio [Sitio Web en Internet] Cumbre del

Milenio. Naciones Unidas, N. Y; 6-8 septiembre 2000. Disponible en: <http://www.un.org/spanish/milleniumgoals/index.html >.

(26) Deutsche Stiftung für internationale Entwicldung (DSE). Digital Inclusion Impact and Challenges of the Networked Economy for Developing Countries. [Website in Internet] . Internationale Welterbildung und Entwiclung gGmblt. February 15, 2001. Available from: <http://www.dse.de/ef/digital/rep0201e.htm>

CAPÍTULO 2

2- El uso de un software libre para la creación de bibliotecas digitales

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH).

Palabras clave: Desarrollo Sostenible, Estrategias de Inclusión Digital. , Sociedad de la Información, Software Libre, UNESCO.

2.1- Introducción

La libertad para usar un programa significa la libertad de cualquier persona u organización de usarlo en cualquier tipo de sistema informático, para cualquier clase de trabajo, y sin tener obligación de comunicárselo al desarrollador o a alguna otra entidad específica (1).

El Software Libre por definición se refiere a la libertad de los usuarios para ejecutar, copiar, distribuir, estudiar, cambiar y mejorar el software. Esto se concreta en cuatro atributos (2):

- La libertad de usar el programa, con cualquier propósito (*libertad 0*).
- La libertad de estudiar cómo funciona el programa, y adaptarlo a tus necesidades (*libertad 1)*. El acceso al código fuente es una condición previa para esto.
- La libertad de distribuir copias, con lo que puedes ayudar a tu vecino (*libertad 2*).
- La libertad de mejorar el programa y hacer públicas las mejoras a los demás, de modo que toda la comunidad se beneficie. (*libertad 3*). El acceso al código fuente es un requisito previo para esto.

2.2- El surgimiento del software libre

El surgimiento del software libre que autorice el uso, la copia y la distribución a cualquiera, ya sea literal o con modificaciones, gratis o mediante una gratificación, (2) es uno de los elementos que caracteriza el momento en que vivimos.

Es necesario destacar que el término software de ``fuente abierta" es usado por algunas personas con la intención de dar a entender más o menos lo mismo que software libre, pero no tiene el mismo significado", el "Software de código fuente abierto" describe una categoría de licencias de software casi, pero no completamente, igual que "software libre" (3). El Software de código fuente abierto incluye software libre, pero también incluye programas semi-libres, e incluso algunos programas privativos bajo su licencia inicial.

La utilización del Software Libre según el criterio y las palabras del propio Stallman precursor del movimiento por el uso del software libre. (4). "*no es un argumento tecnológico, si no que debe plantearse como un modelo diferente de negociación y de producción de conocimiento que tiene*

profundas implicaciones políticas".

El Proyecto GNU (http://www.gnu.org) que lidera la producción de software libre se inició en 1984 con el propósito de desarrollar un sistema operativo compatible con Unix que fuera software libre. Aunque a menudo se haga referencia a estos sistemas como "Linux", la forma correcta de denominarlos es sistemas GNU/Linux. Hoy en día se utiliza una gran variedad de sistemas GNU con un núcleo Linux (2).

El proyecto GNU y la licencia publica general (5) que lo sostiene, establece un régimen jurídico, en que el autor permite toda utilización, modificación y difusión de su trabajo — así como de las versiones derivadas — a condición de que el difusor conceda la misma licencia para los destinatarios — y por lo tanto no restrinja sus derechos.

Algunos especialistas han analizado y reflexionado sobre las razones por las cuales el software libre tiende a ser de alta calidad (6) y estiman que una razón es que por un lado el software libre consigue involucrar a toda la comunidad para que trabaje unida para arreglar los problemas y la otra es que los desarrolladores se preocupan realmente de la fiabilidad ya que un autor que pone el código fuente al alcance de la vista de todos, arriesga su reputación, y por tanto es necesario hacer el software limpio y claro, bajo pena de la desaprobación de la comunidad.

Este movimiento pro utilización del software libre encuentra soporte y desarrollo en organizaciones como la Free Software Foundation (http://www.fsf.org) una organización no lucrativa radicada en Boston, Maryland, Estados Unidos que intenta establecer los vínculos necesarios y apoyar el trabajo de todos aquellos interesados en el uso de software libre y preservar, proteger, y promover la libertad de uso, de estudio, de copia, de modificaciones y de redistribución de software de computadora y los derechos de los usuarios de estos productos.

La Free Software Foundation Europe o FSF Europa (http://www.fsfeurope.org/), fundada posteriormente y que trabaja en todos los aspectos relacionados con el software libre en Europa, y especialmente el Proyecto GNU.

La Fundación Software Libre de América Latina, (http://www.fsfla.org) establecida en la ciudad de Buenos Aires, cuyo objetivo se anuncia como el de promover y defender el uso y desarrollo de software libre en nuestra región, establecida el año pasado.

La *Free Protocols Foundation* (http://www.freeProtocols.org) que se erige como un forum publico independiente, dedicado a apoyar la creación de protocolos libres de patentes ya que la presencia de patentes en un protocolo coloca restricciones en su implementación y uso.

La *League for Programming Freedom* (http://progfree.org) que es una organización que se opone a las patentes de software y a los derechos de autor de las interfaces de usuarios.

En paralelo a este movimiento por el uso del software libre se vienen librando batallas por el acceso sin restricciones a la producción científica, se trata del movimiento *Open Access* el cual existe gracias a la posibilidad de disponer de un software libre y la documentación necesaria para su puesta en marcha.

2.3- La informatización de la Sociedad Cubana

La sociedad cubana se encuentra enfrascada en un proceso de informatización con un programa rector concebido desde 1997, que incluye aspectos tales como:

- 1- Desarrollo de la Industria Cubana de las Tecnologías de la Información.

- 2- Servicios Integrales para los ciudadanos.
- 3-- Informatización del Gobierno, la Administración y la Economía.
- 4- -Informatización Territorial (7).

En consonancia con las orientaciones de los organismos internacionales responsabilizados con la conducción del proceso de implementación de una Sociedad de la Información/Conocimiento, nuestro país ha emprendido una serie de acciones para lograr este empeño.

En la Declaración de Principios de la Cumbre Mundial sobre la Sociedad de la Información celebrada en diciembre de 2003 (8) queda muy claramente expresado entre otros aspectos: "*La capacidad universal de acceder y contribuir a la información, las ideas y el conocimiento es un elemento indispensable en una Sociedad de la Información integradora*".

Desde diciembre de 1999 comenzó la utilización masiva de medios audiovisuales en las escuelas cubanas, lo cual generó la necesidad de producir programas de televisión destinados a los centros escolares. Se ha puesto en funcionamiento un nuevo canal televisivo dedicado esencialmente a la educación, que logra una cobertura de más del 85 % de la población y transmite más de 15 horas diarias como promedio.

Las universidades cubanas están conectadas a Internet. Así mismo ocurre con los centros científicos, órganos de prensa y otras instituciones que han sido priorizadas en la creación de las condiciones técnicas a que se ha podido llegar, no obstante los obstáculos propios de una nación subdesarrollada, bloqueada y con pocos recursos financieros.

Además de los programas y proyectos mencionados, el programa rector de informatización, concebido desde 1997, incluye otros aspectos tales como:

- Desarrollo de la Industria Cubana de las Tecnologías de la Información (Acuerdo No. 3736 de fecha 18 de julio del 2000, adoptado por el Comité Ejecutivo del Consejo de Ministros de Cuba)
- Servicios Integrales para los ciudadanos (Ministerio de la Informática y las Comunicaciones de Cuba (http://www.mic.gov.cu/hinfosoc.aspx)
- Informatización del Gobierno, la Administración y la Economía (Ministerio de la Informática y las Comunicaciones de Cuba (http://www.mic.gov.cu/hregulacion.aspx)
- Informatización Territorial (Portal del Ciudadano Cubano. Referencia para Trámites y Servicios en nuestro país (http://ciudadano.infosoc.cu/ciudadano/)

Cuba continuará la informatización como parte de su lucha por elevar la calidad de vida del pueblo cubano y lograr una sociedad cada vez más justa, equitativa y solidaria. Un mundo mejor es posible.

2.4- La utilización de Greenstone en la creación de bibliotecas digitales en Cuba

En la reunión de UNESCO "*Towards Knowledge Societies*" realizada en el marco de la 32ª sesión de la Asamblea General en 2003, se establecieron los postulados que debían guiar la construcción de una Sociedad del Conocimiento: libertad de expresión, acceso universal a la información y al conocimiento, respeto por la dignidad humana y la diversidad cultural y lingüística, calidad de la educación para todos, inversiones en ciencia y tecnología y comprensión e inclusión de los sistemas de conocimientos nativos (9).

En el Manifiesto de Alejandría: la Sociedad de la Información en acción (10), la IFLA (*International Federation of Library Associations and Institutions*) urge a los gobiernos nacionales,

regionales y locales y a los organismos internacionales (entre otras cosas) a que:

- inviertan en las bibliotecas y servicios de información como elementos vitales dentro de sus estrategias, políticas y presupuestos para la Sociedad de la Información;
- mejoren y extiendan las redes bibliotecarias ya existentes para el mayor beneficio posible de sus ciudadanos y comunidades;

Las políticas de nuestro Estado en perfecta consonancia con los postulados de los organismos internacionales encargados de guiar la implementación de una Sociedad de la Información/ Conocimiento ha emprendido las tareas necesarias para lograr la informatización de la sociedad y ha iniciado las acciones necesarias para sumarse al Desarrollo de la Red Latinoamericana de soporte para Greenstone (Figura 2.1) , para lo cual este libro constituirá un pilar importante.

Este software es el fruto del trabajo del Proyecto Biblioteca Digital de Nueva Zelanda y es distribuido por la Unesco y la ONG en Información para el Desarrollo Humano bajo los términos de la licencia Pública General de GNU (http://www.gnu.org/gnu/gnu-history.es.html).



Figura 2.1-Red Latinoamericana de soporte a Greenstone

El software libre Greenstone es un completo sistema diseñado para construir y organizar colecciones de recursos digitales, creado por el Proyecto de Biblioteca Digital de Nueva Zelanda con sede en la Universidad de Waikato (11) , y que es desarrollado y distribuido por la UNESCO con la colaboración de la ONG "Información para el Desarrollo Humano" con sede en Bélgica

El proyecto de desarrollo de la Red Latinoamericana de soporte para Greenstone, se propone crear centros nacionales en todos los países del área y ya ha iniciado sus actividades con la denominación de tres centros regionales en su inicio, situados en Argentina, Chile y Cuba. Se ha orientado además la creación de una red de usuarios de Greenstone liderada por estos centros regionales y soportada en http://www.greenstone-la.org

2.5- Consideraciones finales

El establecimiento de Redes de Colaboración y/o de Conocimientos puede constituir una estrategia alcanzable para los países de economías menos desarrolladas donde cada uno aporte de acuerdo a

sus posibilidades y reciba de acuerdo a sus necesidades.

2.6- Referencias Bibliográficas

(1) Stallman R. Categorías de Software Libre y No Libre. [Sitio Web en Internet]. Proyecto GNU. Free Software Foundation. 23 May 1999. Disponible en: <http://www.gnu.org/philosophy/categories.es.html>.

(2) Free Software Foundation. La Definición de Software Libre. [Sitio Web en Internet]. Proyecto GNU. Free Software Foundation. 2006/ 05/20. Disponible en: <http://www.gnu.org/philosophy/free-sw.es.html >.

(3) Stallman R. Porqué "Software Libre" es mejor que software de "Código Fuente Abierto". [Sitio Web en Internet]. Free Software Foundation. 6 de noviembre de 1999 Disponible en: <http://www.gnu.org/philosophy\free-software-for-freedom.html>.

(4) Stallman R. Free Software as a Social Movement. Justin Podur interviews Richard Stallman. Zmagazine. ZNet Top. [Seriada en Internet]. December 18, 2005. Available from: <http://www.zmag.org/content/showarticle.cfm>.

(5) GNU .Filosofía del Proyecto GNU. [Sitio Web en Internet]. 2006/05/20. Disponible en: <http://www.gnu.org/philosophy/philosophy.es.html>.

(6) Free Software Foundation. ¡El Software Libre es Más Confiable! [Sitio Web en Internet]. Proyecto GNU - Free Software Foundation (FSF) 2 de Mayo, 1999). Disponible en: <http://www.gnu.org/software/reliability.es.html>.

(7) Del Brutto, Bibiana Apolonia y otros. "La Sociedad del Conocimiento, presentación oficial de la República de Cuba ante la Cumbre de la Sociedad de la Información en Ginebra". ARCHIVO del Observatorio para la CiberSociedad , [Sitio Web en Internet]. 2003. Disponible en: < http://www.cibersociedad.net/archivo/articulo.php?art=155>.

(8) CMSI. Construir la Sociedad de la Información: un desafío global para el nuevo milenio. [Sitio Web en Internet]. Declaración de Principios. Documento WSIS-03/GENEVA/4-S.12 de mayo de 2004. Disponible en: <http://www.itu.int/wsis>.

(9) UNESCO: Towards Knowledge Societies. Round Table Meeting. 32nd session of UNESCO's General Conference, [Website en Internet]. Paris, October 9 and 10, 2003. Available from:<http://portal.unesco.org/ci/en/ev.php-URL_ID=12420&URL_DO=DO_TOPIC&URL_SECTION=201.html>.

(10) IFLA. Manifiesto de Alejandría sobre Bibliotecas: la Sociedad de la Información en Acción. [Sitio Web en Internet]. 24-02-2004 Disponible en: <http://www.ifla.org/faife/policy/iflastat/gldeclar-e.html>

(11) Witten, Ian H. and David Bainbridge. A brief history of the Greenstone Digital library Software. [Website on Internet]. 2007-01. Available from: <http://wiki.greenstone.org/wiki/gsdoc/others/Greenstone_history.htm>

CAPÍTULO 3

3- Características de las Bibliotecas Digitales en el Siglo XXI

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH)

Palabras clave: Alfabetización Digital, Alfabetización en Información, Bibliotecas, Bibliotecas Digitales, Desarrollo Sostenible, Economía Digital, Estrategias de Inclusión Digital. , Sociedad de la Información.

3.1- Introducción

Las Bibliotecas son entidades que se encargan de procesar información como tarea básica. El valor social y económico de la información en la era de la información y el conocimiento, proviene del reconocimiento individual o colectivo que se le confiere, de la necesidad de la sociedad por utilizarla o poseerla (1)

Según P. Drucker (2), un importante Gurú de la Gerencia de la Información /Conocimiento; la información interna en una institución se convierte en un recurso competitivo pues en la sociedad actual el conocimiento es un recurso principal para la actividad económica, ya que las instituciones deben explotar continuamente nuevas oportunidades de negocio con el fin de crear potenciales fuentes de crecimiento y por su naturaleza intangible y sus características de codificabilidad y especificidad se convierte en un recurso escaso y valioso que proporciona la base para desarrollar estrategias que superen las de los competidores. (2). Esta idea nos convence de que el procesamiento de la información resulta un elemento clave en la sociedad actual.

En la implementación y desarrollo de una Sociedad de la Información /Conocimiento resulta imprescindible utilizar inteligentemente este valioso recurso y preservarlo. Son pues las bibliotecas, las instituciones destinadas a este fin que pueden lograr este propósito de forma efectiva y conveniente.

Estimamos conveniente ahora proveer en primer lugar una clasificación sencilla de lo que representa el concepto de biblioteca y utilizaremos la propuesta por Sharon y Frank en 2008 (3):

- Analógicas/Bibliotecas de Papel (BP) - la biblioteca de papel clásica con su fichero manual.
- Automatizada/Biblioteca Híbrida (BH) - una biblioteca analógica con el catálogo automatizado.
- Biblioteca Digital (BD) - una biblioteca automatizada en la que la mayor parte de la información es digital.

3.2- Evolución histórica y tecnológica de los servicios bibliotecarios

Se impone efectuar una caracterización de los diferentes modelos o formas que han adoptado las

bibliotecas desde su concepción original.

Existe una abundante bibliografía al respecto en la que se pretende catalogar las bibliotecas de acuerdo a los atributos que presentan en cualquiera de sus modalidades.

El concepto de Biblioteca tradicional (4) según la ALA (American Library Asociation) define la biblioteca como una “Colección de material de información organizada para que pueda acceder a ella un grupo de usuarios. Tiene personal encargado de los servicios y programas relacionados con las necesidades de información de los lectores”. Esta idea primaria ha ido evolucionando hacia conceptos mas elaborados.

Algunas de estas ideas son las siguientes (5):

- Biblioteca tradicional. Espacio físico y dinámico en el que se ofrecen servicios y procesos manuales.
- Biblioteca electrónica. Espacio físico en el que se han automatizado algunos servicios, como catálogos y procesos técnicos.
- Biblioteca virtual. Modelo de biblioteca que se construye actualmente. Consiste en una biblioteca sin muros y sin espacios físicos, donde los documentos se localizan electrónicamente, y el acceso a los mismos tiene lugar mediante la virtualidad.
- Biblioteca híbrida. Es una combinación de la biblioteca electrónica y la virtual. Oferta servicios in situ y virtuales. Posee, tanto acervo impreso como electrónico. Cuenta con puntos de acceso virtuales que posibilitan al usuario conocer como obtener la información.

3.3- Surgimiento de las bibliotecas digitales

Consideremos primero un criterio autorizado de lo que representa una biblioteca digital como proyecto y asumamos la definición hecha por la Federación de Bibliotecas Digitales en el año 2003 (6):

> *"Las bibliotecas digitales son organizaciones que fomentan los recursos, incluido el personal especializado, para seleccionar, estructurar, ofrecer acceso intelectual, traducir, distribuir, preservar la integridad y garantizar la permanencia de las colecciones digitales, de forma que estén disponibles para una o varias comunidades en un ambiente internacional coordinado"*

Existen otras definiciones que estimamos debemos reflejar aquí. Así por ejemplo en el *Workshop on Distributed Knowledge Work Environments* sobre Bibliotecas Digitales efectuado en Santa Fe, se postuló lo siguiente:

> *"entorno donde se reúnen colecciones, servicios, y personal que favorece el ciclo completo de la creación, difusión, uso y preservación de los datos, para la información y el conocimiento (7)"*

Otra definición interesante aporta la *Association of Research Libraries* (ARL):

> *"repertorios de objetos digitales, más o menos organizados, que sirven a una comunidad de usuarios definida, los cuales tienen los derechos de autor presentes y gestionados, y disponen de mecanismos de preservación y conservación. Esta definición tiene en cuenta que estos repertorios constan de datos (el contenido) y metadatos (la información que describe los datos) e incorporan técnicas de busca y*

recuperación de la información (8)".

Greenstone es una herramienta constituida por un grupo de programas para construir bibliotecas digitales (9). Ha sido desarrollado y distribuido con la cooperación de la UNESCO. Este software es el fruto del trabajo del Proyecto Biblioteca Digital de Nueva Zelanda y es distribuido por la UNESCO y la ONG en Información para el Desarrollo Humano bajo los términos de la licencia Pública General de GNU ((http://www.gnu.org/gnu/gnu-history.es.html).

El Proyecto de la Biblioteca Digital de Nueva Zelanda se inicio en 1995, comenzando con una colección de 50,000 reportes técnicos descargados de Internet (10).

No fue sino hasta finales de 1997, que fue adoptado el termino Greenstone. Que es una forma de piedra de jade muy valorada por los nativos Māori. La decisión de colocar el software como una herramienta de código abierto y de utilizar la licencia Publica General de GNU, se produjo alrededor de este momento. (10)

Por su fiabilidad, nivel de desarrollo y de mantenimiento, la Unesco lo ha incluido en su programa de aplicaciones informáticas para servicios de información y documentación. Igualmente, por sus prestaciones a la comunidad internacional se hizo acreedor del premio otorgado por la fundación Andrew Mellon durante 2010 la cual reconoció, de este modo, su aporte al ámbito de la educación, la cultura y la promoción humana. (Proyecto Greenstone http://www.greenstone.org).

3.4- Ventajas y Desventajas del uso de Bibliotecas Digitales

3.4.1- Efectividad y desempeño de las Bibliotecas Digitales

Los cambios tecnológicos han redireccionado la idea de las bibliotecas para referirse a ellas como lugares destinados a la salvaguarda de la memoria de los pueblos, las civilizaciones y las culturas, así como también a repensar el proceso y el acceso a la información (11). Consideramos que esta idea encierra una compleja situación de orden tecnológico, social y político que debemos enfrentar.

No se trata tan solo del hecho de digitalizar los recursos de información y colocarlos en tecnología de redes convenientemente. Este fenómeno encierra otras dificultades tales como el denominado copyright o derecho de autor, que en la época digital en que vivimos adquiere una connotación importante y da lugar al movimiento Copyleft de Richard Sallman (12) que ya hemos mencionado antes y a iniciativas de acceso abierto a la información como las que hemos citado en el capitulo 2 de este libro.

Pudiéramos establecer un paralelo entre las desventajas de una biblioteca tradicional y las ventajas que pueden ofrecernos las bibliotecas virtuales, destacando algunas de las pertinentes reflexiones que efectúa Mojica L. (13):

- Desventajas de una biblioteca tradicional:
  - Escasez de servicios frente a la virtual, limitándose, en general, a la consulta y lectura de documentos en sala, el préstamo a domicilio, un servicio de préstamo interbibliotecario limitado por la dificultad de acceso a los catálogos de otras bibliotecas, servicios de referencia y asesoramiento, etc.
  - Limitación en el uso debido al horario establecido en la biblioteca tradicional.
  - Fuentes de información no actualizadas generalmente.
  - Personal no es suficiente para satisfacer las necesidades de los usuarios.
  - Servicios casi exclusivamente locales, determinados por la distancia geográfica..

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

– Grandes limitaciones a la cooperación bibliotecario por la inexistencia de catálogos en línea, que imposibilita la adquisición compartida o cooperativa efectiva.

- Ventajas de una biblioteca virtual:
  – -La biblioteca virtual provee para ofrecer todo tipo de información bibliográfica, de manera eficaz, y sobre todo, actualizada en el mismo momento en que la generamos, y esto se puede hacer por diferentes vías: por autores, por materias.
  – -La biblioteca virtual cuenta con la generación de ayudas en línea, necesarias para los usuarios.
  – -La biblioteca virtual provee para la comunicación efectiva y rápida mediante el uso de correos electrónicos.

Pero en el caso de las Bibliotecas Virtuales, lamentablemente todos los atributos no representan ventajas y así tenemos el caso de la utilización de Internet y la utilización de los recursos de información disponibles allí , con un orden caótico en su colocación y además exento de los controles de calidad que brindan las editoriales y comités revisores de revistas.

Tenemos que considerar además como una desventaja de las bibliotecas virtuales el costo de acceso (14).

En síntesis, el usuario virtual debe conocer las fuentes y todos los recursos y herramientas que precisa, las tecnologías de la información y de la investigación, logrando así una mayor movilidad e independencia, y otras oportunidades de búsqueda en la red (15), pero también asume el compromiso insoslayable de convertirse en su propio rastreador de los recursos de información que precisa para satisfacer sus necesidades informativas y de comprobar la pertinencia y calidad de los mismos, lo cual puede ser salvado con una biblioteca digital de libre acceso que cuente con colecciones bien elaboradas.

3.5- Herramientas para la construcción de Bibliotecas Digitales de libre acceso

Se destacan actualmente en la literatura tres programas o herramientas que merecen nuestra consideración en el caso de que intentemos construir una biblioteca digital.

3.5.1- DSpace

DSpace (http://www.dspace.org/) es un software de código abierto diseñado por el *Massachusetts Institute of Technology* (MIT) y los laboratorios de HP para gestionar repositorios de ficheros (textuales, audio, vídeo, etc.), facilitando su depósito, organizándolos en comunidades, asignándoles metadatos y permitiendo su difusión a recolectores o agregadores (16).

DSpace se una aplicación cliente / servidor que se gestiona vía Web, es decir, que la mayor parte de procesos pueden llevarse a cabo con un navegador estándar como Internet Explorer, Firefox u Opera. Desde el punto de vista del servidor, DSpace puede instalarse tanto en entornos Linux como en un servidor Windows, siempre que reúna una serie de pre-rrequisitos (16).

DSpace utiliza el estándar de metadatos Dublin Core (DCMI) para la descripción de los documentos, que van desde texto, hasta imágenes y videos, entre otros, y que luego posibilita su recuperación precisa. La interfaz en ambiente Web varía en dependencia de la persona que la utilice; así, los responsables de la colección tienen una, los administradores del sistema otra y los usuarios finales otra; aunque todas coinciden en que son en idioma inglés. Los usuarios pueden hacer sus búsquedas a partir de los metadatos declarados por los especialistas o simplemente

mediante las listas de documentos por campos específicos, como son: autor, título y fecha (17).

La DCMI es la iniciativa internacional de metadatos más sólida e importante para la organización y recuperación de información en Internet en forma normalizada, eficaz y con un propósito general; además, el conjunto de elementos Dublin Core es un estándar internacional ISO 15836-2003, con independencia de que, en cada ámbito de información, se desarrollen esquemas propios de meta-información para mejorar la recuperación en la red (17).

3.5.2- Fedora

Disponible en: (http://www.fedora.info/). Requiere la instalación de Java, pero también permite trabajar con varias opciones de bases de datos, entre ellas MySQL, PostgreSQL y McKoi SQL, incluyendo esta última para facilitar la instalación de entornos de prueba y desarrollo sin necesidad de configurar un sistema complejo de bases de datos (18).

En adición, también requiere Apache Tomcat, pero el instalador de Fedora incluye tanto la base de datos McKoi SQL como el Tomcat, facilitando el proceso. De hecho para habilitar e instalar Fedora, sólo hace falta tener Apache corriendo en un servidor Windows o Linux, instalar Java *Runtime Environment* y descargar e instalar el *Fedora Installer*. La desventaja de esta herramienta radica es que resulta necesario modificar manualmente una serie de variables y valores en los archivos de configuración, lo cual torna algo compleja su utilización (18)

3.5.3- Greenstone

Greenstone (http://www.greenstone.org/) consta, de un motor de indización y recuperación de la información textual: MG que utiliza el modelo vectorial para el tratamiento de la información. Este es alimentado por un conjunto de scripts en Perl, los que pre procesan los documentos mediante los plugins o pequeños programas que transforman los documentos originales en documentos XML, codificados en Unicode (19). Esto asegura la capacidad de la aplicación para tratar cualquier idioma.Greenstone puede ejecutarse en la mayoría de los Sistemas Operativos. Viene preinstalado en las diferentes distribuciones de Linux y está integrado para las distintas versiones de Windows y Macintosh (20)

Se comporta mejor en una PC con plataforma Linux que con Windows pero su fácil instalación en servidores hace que se desempeñe correctamente con servidores que combinan software propietario y herramientas de software libre.

3.6- Etapas de la Actividad Científico Informativa en Greenstone

Las etapas de la Actividad Científico - Informativa que se describen son las siguientes (4):

3.6.1- Recolección

Es necesario recolectar, con la mayor exhaustividad posible, la información necesaria. Debe seleccionarse y actualizarse regularmente la información útil, sólida y significativa para el trabajo científico o práctico de una o varias disciplinas, en una tendencia de investigación o rama del conocimiento, entre la colosal cantidad de información existente (21).

Desde la versión 2.41 se ha incorporado una interfaz gráfica en Java, GLI o *Greenstone Librarian Interface*, que facilita el proceso de creación y administración de las colecciones (21).

Esta aplicación en Java, denominada Interfaz del Bibliotecario, brinda opciones de configuración diseñadas para facilitar el trabajo de personalización y se ejecuta tanto en modo local como con un

servidor Web (22) A continuación se presenta una imagen que muestra su apariencia.

Figura 3.1- Interfaz de trabajo del bibliotecario del software Greenstone

3.6.2- Procesamiento analítico – sintético

Efectuar un procesamiento analítico-sintético representa incorporar cada grupo de datos registrados en un documento y analizarlo desde el punto de vista de su contenido. Los resultados de dichos análisis se anotan o sintetizan brevemente con ayuda de una notación especial, desarrollada a los efectos, y basada en palabras del lenguaje natural, letras, cifras, fórmulas, tablas, diagramas. La información así procesada, se torna adecuada para los subsiguientes pasos: almacenamiento, recuperación y uso (4).

En el caso de Greenstone. el proceso básico incluye la selección de los documentos a tratar, su procesamiento y conversión en XML mediante los plugins de Perl adecuados, la extracción de metadatos de los documentos (los que fueron asignados manual o automáticamente previamente), la indización con el motor elegido y el almacenamiento con GDBM (GNU DataBase Manager) (19).

Cuenta también Greenstone con algoritmos de indexamiento únicos (23) conocidos como Hash (función o método para generar claves o llaves que representen de manera univoca a un documento, registro, archivo, resumir o identificar un dato a través de la probabilidad, utilizando una función hash o algoritmo hash) que no clasifica como un índice, pero proporciona una fórmula que nos crea una dirección para un registro y una estructura "trie" que permite ubicar más fácilmente la dirección entre bloques de datos (24).

Es preciso apuntar que el procesamiento y análisis de documentos XML están basadas en 2 interfaces de aplicación SAX (Simple Api for XML) y DOM (Document Object Model) las cuales analizan el documento y son capaces de construir su árbol de objetos (elementos, atributos, etc.) para poder realizar búsquedas o transformaciones en él. La Figura 3.2 (24) muestra un ejemplo ofrecido por este autor del procedimiento que realizan las hojas de estilo en XML para convertir un documento XML en uno HTML, analizando el documento y modificando partes específicas de un lenguaje al otro.

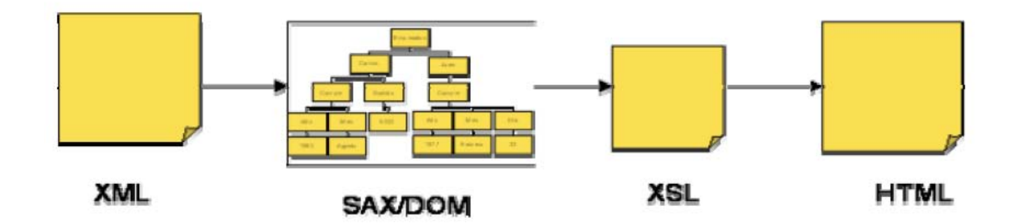


Figura 3.2 Conversión de documentos XML a HTML

3.6.3- Almacenamiento

El objetivo de este proceso es garantizar que la información seleccionada ocupe el menor espacio posible, que los medios de registro ofrezcan un largo período de almacenamiento y que puedan disponerse en un sistema ordenado que permita la rápida y fácil recuperación de los datos relevantes (4).

El tratamiento y almacenamiento de la información en Greenstone consta de otro software libre del ul ya hemos hablado: el sistema de gestión de bases de datos GDBM (GNU DataBase Manager) disponible en: http://www.gnu.org/software/gdbm/gdbm.html.

El paquete incluye una versión especial de MG, el MG+/MGPP, para cuando fueran necesarios formularios con múltiples campos para la interfase de búsqueda. Desde la versión 2.61 se incluye el motor Lucene, de la Fundación Apache (19)

3.6.4- Recuperación de la información

La recuperación consiste en una serie de operaciones lógicas que garantizan la localización de la información que se necesita. Cualquier procedimiento empleado para desarrollar esta tarea debe permitir, en cualquier momento, recuperar de una colección, aquella información que proporcione una respuesta directa a la pregunta formulada por el usuario (4).

El Greenstone implementa el protocolo CORBA y esto permite trabajar tanto con direcciones locales como distribuidas (25).

El Greenstone consta, de un motor de indización y recuperación de la información textual: el MG que ya hemos mencionado que utiliza el modelo vectorial para el tratamiento de la información. Este es alimentado por un conjunto de scripts en Perl, los que preprocesan los documentos mediante los plugins o pequeños programas que transforman los documentos originales en documentos XML, codificados en Unicode. Esto asegura la capacidad de la aplicación para tratar cualquier idioma (19).

3.6.5- Diseminación

La diseminación como proceso de la actividad científico-informativa nació como consecuencia de que los resultados de los procesos anteriores y la recuperación de la información, pueden presentar simultáneamente interés capital para un elevado número de usuarios. Los documentos (o sus partes) con información relevante deben reproducirse de modo que se obtenga un número suficiente de copias idénticas, con vistas a su rápida diseminación entre los interesados (4).

Elemento clave en la diseminación y uso en Greenstone lo constituye las características del fenómeno de interoperabilidad con que cuenta el sistema:

- La interoperabilidad se manifiesta fundamentalmente por (26):
- La capacidad de los sistemas para trabajar entre sí en tiempo real o programado.
- La capacidad del software para trabajar en diferentes sistemas
- La capacidad de los datos para ser intercambiados entre diferentes sistemas (portabilidad).

La interoperabilidad puede ser clasificada en: Interoperabilidad Sintáctica, Interoperabilidad Semántica e Interoperabilidad Estructural (27), Esta última adquiere un connotación especial en Greenstone pues la comunicación e interactuación en ambientes heterogéneos brinda una gran fortaleza al sistema Greenstone al poder utilizar protocolos especializados como Z39.50, OAI-PMH, SRU,RSS, y otros.

Este aspecto es tan importante que será tratado nuevamente en otro capitulo de este libro.

3.6.6- Uso de la información

Las actividades científico-informativas están en función del uso que se haga de los resultados. Por esta razón, entre los objetivos de estas actividades se incluyen: la popularización, tanto de los logros de la ciencia y la tecnología como de las técnicas más eficientes para el manejo de los documentos, así como el estudio de los resultados que genera el uso de dichos documentos y el impacto causado por el trabajo de información en la ciencia, la técnica, la economía, la educación, la cultura, etc. (4).

La Organización Internacional de Normalización en su norma ISO 9241 define la usabilidad como:

> *"la medida en la que un producto puede utilizarse por sus usuarios específicos para conseguir ciertos objetivos particulares con efectividad, eficiencia y satisfacción en un contexto de uso concreto" (27).*

Si entre los Atributos de Usabilidad contamos con (27):

- Facilidad de aprendizaje,
- Eficiencia,
- Recuerdo en el tiempo,
- Satisfacción y Tasa de errores negativa.

Nos percatamos que todos ellos son atributos que encontramos en el sistema Greenstone.

3.7- Consideraciones Finales

Con el Sistema Greenstone Ud. desarrolla fácilmente todas las actividades que caracterizan la actividad científico-informativa de modo automático y sin el mayor esfuerzo con procedimientos que son transparentes para Ud. y que hemos querido destacar en el desarrollo de este capítulo

3.8- Referencias Bibliográficas

(1) Morales E. La Sociedad de la Información en el Siglo XXI y la Biblioteca Universitaria. Revista Digital Universitaria [Seriada en Internet] 30 de Junio de 2001 Vol.2 No.2. Disponible en:<http://www.revista.unam.mx/vol.2/num2/art1/index.htm>

(2) Drucker P. La Organización Basada en la Información. Barcelona. Grupo Editorial Norma.1996

(3) Sharon T Frank A. J. Bibliotecas digitales en Internet - 66th IFLA Council and General Conference [Sitio Web en Internet] Conference Programme and Proceedings 2000-08-22 Disponible en:< http://www.ifla.org/new website>

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

(4) ALA. American Library Association. Handbook of organiztion [Book in Internet]. 2009/09/15. Available from: <http://www.ala.org/aboutala/governance/handbook/index.cfm >

(5) González C. Desarrollo de servicios digitales en las bibliotecas: nuevos retos y nuevos escenarios. En: Jornadas Nacionales de Bibliotecas Universitarias. Universidad Privada Antenor Orrego, Perú. [Sitio Web en Internet]. 2005. Disponible en: <http://eprints.rclis.org/archive/00007906/01/Desarrollo_de_servicios_digitales_en_las_bibliotecas__-_20%E2%80%A6.pdf>

(6) York V. A biblioteca digital. Recursos e projetos. [Sitio Web en Internet].2003 Disponible en: < http://www.cg.org.br/gt/gtbv/YorkatIBICT/tsld004.htm>

(7) Santa Fe Workshop on Distributed Knowledge Work Environments. Digital Libraries [Website in Internet]. -March 9-11,1997. Available from: <http://www.dli2.nsf.gov/>

(8) Association of Research Libraries. "Definition and Purposes of a Digital Library". [Website in Internet].October 23, 1995. Available from: <http://sunsite.berkekey.edu/ARL/definition.htm/>

(9) Witten I, Boddie S. Thompson D. Guía del usuario. Nueva Zelanda: Universidad de Waikato [Sitio Web en Internet], Departamento de Informática, 2005. Disponible en: <http://www.greenstone.org/manuals/gsdl2/style.css>

(10) Witten I, Bainbridge D. A brief history of the Greenstone Digital Library Software.[Web site in Internet] University of Waikato, Hamilton, New Zealand. 2007-01-15. Available from: <http://wiki.greenstone.org/wiki/gsdoc/others/Greenstone_history.htm#_ftnref2>

(11) Anglada de Ferrer L. Biblioteca Digital: ¿mejor, peor o sólo distinto? Anales de la Documentación. [Seriada en Internet]. 2008-02-07. Disponible en: <http://revistas.um.es/analesdoc/article/view/2521>

(12) Stallman M. ¿Qué es el copyleft? [Libro en Internet] Capitulo 14 Software libre para una sociedad libre. Versión 1.0 Primera edición en castellano (papel): Edición electrónica a cargo de: Miquel Vidal. Edición: Traficantes de Sueños Madrid. Noviembre 2004.Disponible en: <http://sindominio.net/traficantes>.

(13) Mojica L. La Efectividad de la Biblioteca Virtual Vs. [Sitio Web en Internet] Slidesharecdn.com. 2011. Disponible en: <http://public.slidesharecdn.com/v3/styles/combined.css?26846>.

(14) Babini D. Fraga J. Alcances del concepto de Biblioteca Virtual. Cap. 2. En: Bibliotecas Virtuales para las Ciencias Sociales. [Libro en Internet]. 2006. Disponible en: http://bibliotecavirtual.claso.org.ar/libros/biblio/cap2.pdf

(15) Clavero, J., Codina, M. Pérez, A. Serrat, M, Estudio de caso de servicio de préstamo de libros electrónicos. El Profesional de la Información [Seriada en Internet]. 2009, no. 2; p. 237-241 Disponible en: <http://hdl.handle.net/10760/14662>.

(16) Rodríguez-Gairín J.M, Sulé A. DSpace: un manual específico para gestores de la información y la documentación.[Libro en Internet] junio de 2008. Disponible en: <http://www.ub.edu/bid/20rodri2.htm>

(17) Sarduy Domínguez Y, Urra González P. Herramientas para la creación de colecciones digitales. ACIMED [Seriada en Internet] 2006;14 (5). Disponible en: <http://bvs.sld.cu/revistas/aci/vol14_5_06/aci19506.htm>

(18) Miranda C. Sistemas de Gestión de Documentos bajo Licencia de Código Libre [Sitio Web en

Internet] 08 Junio 2009. Disponible en: <http://www.educar.org//sites/socinfo.com/files/favicon.ico>

(19) Bordenave M.G., Fernández M., Sosa O La arquitectura de Greenstone. En: Software libre de publicación y difusión de colecciones de publicaciones científicas. [Libro en Internet] Análisis de Greenstone. Grupo CEDIPUR. Instituto de Planeamiento Urbano y Regional.Facultad de Arquitectura y Urbanismo – UNNE. 2007. Disponible en: http://arq.unnc.edu.ar/centros/institutos/ipuv/publipuv/greenstone.pdf

(20) Greenstone .org. Biblioteca Digital Greenstone. Guía del Programador .Configuración del sitio Greenstone. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm>

(21) Cabrera, A. M., Coutin A. Las bibliotecas digitales. Parte II. Componentes. Etapas de la Actividad Científico Informativa. ACIMED [Seriada en Internet] 2005; 13(3) : Disponible en: <http://bvs.sld.cu/revistas/aci/vol_13_3_05/aci05305.htm>

(22) Tramullas J. Biblioteca Digital Greenstone. Capitulo 10. En.Software Libre para Servicios de Información [Libro en Internet] Digital. Madrid: Pearson Prentiss Hall, 2006. Disponible en:<http://tramullas.com>

(23) Greenstone.org. Guía de Instalación Greenstone. [Sitio Web en Internet]. 2002. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone>

(24) Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3. El Sistema de Ejecución de Greenstone [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/style.css>

(25) Blanco J. Notas Almacenamiento y Recuperación de Información. En: Lecturas en Ciencias de la Computación [Libro en Internet] RT 2007-08. Centro IOMMA. Caracas, Septiembre, 2007. Disponible en: <http://www.ciens.ucv.ve/jblanco/>

(25) Travieso M. Las publicaciones electrónicas: una revolución en el siglo XXI. [Seriada en Internet] ACIMED v.11 n.2 ene.-abr. 2003. Disponible en: <http://scielo.sld.cu/scielo.php?script=sci_issuetoc&pid=1024-943520030002&lng=es&nrm=iso>

(26) Mari-Carmen Marcos y colaboradores.. Evaluación de la usabilidad en sistemas de información terminológicos online [Seriada en Internet] . "Hipertext.net", núm. 4, 2006. Disponible en: <<http://www.hipertext.net>

(27)- Nielsen J. Usability Engineering. AP Professional, 1993. Citado por: Mari-Carmen Marcos et al.. Evaluación de la usabilidad en sistemas de información terminológicos online [Seriada en Internet]. "Hipertext.net", núm. 4, 2006 Disponible en: <http://www.hipertext.net>

CAPÍTULO 4

4- Los atributos y condiciones de trabajo del software Greenstone

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH)

Ms C Ricardo Fernández Rodríguez (ricardo@cecam.sld.cu) Profesor Auxiliar del I.S.C.M.H. Máster en Informática en Salud. Centro de Cibernética Aplicada a la Medicina (CECAM) I.S.C.M.H. Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba. CP: 11600, Telf.: (537) 271-1354

Palabras clave: Bibliotecas Digitales, Desarrollo de software. Desarrollo Sostenible, Economía Digital, Estrategias de Inclusión Software Libre, UNESCO.

4.1- Introducción

Greenstone, es un completo sistema diseñado para construir y organizar colecciones de recursos digitales, creado por el Proyecto de Biblioteca Digital de Nueva Zelanda con sede en la Universidad de Waikato , es desarrollado y distribuido por la UNESCO con la colaboración de la ONG "Información para el Desarrollo Humano" con sede en Bélgica

Greenstone es un software open source (o software de código abierto) que se distribuye de acuerdo a los términos de la Licencia Publica General GNU (http://www.gnu.org/gnu/gnu-history.es.html)

Existen sendas versiones de Greenstone disponibles para Windows y Unix, como binarios y en forma de código fuente (1).

4.2- Definiciones del sistema Greenstone

Las versiones mas recientes del software Greenstone cuentan con lo que se ha dado en llamar "*la interfaz del bibliotecario*" que no es mas que una aplicación java, de entorno grafico, que se incorpora a partir de la versión 2.41.Esta interfaz gráfica, desarrolla con detalle las opciones de configuración y facilita el trabajo de configuración de la colección que se creará pudiendo ser ejecutada de forma local (2). En las últimas versiones de Greenstone se ha desarrollado un *applet* de java que realiza las mismas funciones a través del servidor Web lo que permite el trabajo remoto (2).

Greenstone utiliza el Sistema Corba (*Common Object Request Broker Architecture*) que proporciona la infraestructura para la arquitectura OMA (*Object Management Architecture*) lo cual permite trabajar objetos en entornos heterogéneos, (los entornos heterogéneos son aquellos en los que las arquitecturas que constituyen el entorno pueden ser sistemas Microsoft Windows, máquinas Unix de diferentes fabricantes Linux por ejemplo, e incluso sistemas como MacOS o OS/, así como protocolos de comunicación como TCP/IP, IPX) (3).

CORBA establece estándares para la comunicación de objetos a través de procedimientos/métodos

remotos, utilizando el lenguaje IDL ("*Interface Definition Language*"), que como su nombre indica define las interfases o estructuras que serán utilizadas en este ambiente CORBA (4).Como veremos mas adelante esta cualidad nos permitirá crear un servidor remoto para la utilización de este software.

Este paradigma del sistema CORBA orientado a objetos, permitirá que Greenstone pueda ser ejecutado en distintas plataformas informáticas (como hemos comentado previamente) e implementarse con diferentes lenguajes de programación, así como también, acceder a la misma serie de objetos distribuidos por Internet (o cualquier otra red). Este sistema CORBA, permitirá además acceder a un servidor remoto de colecciones de Greenstone, escrito en C++ (5).

4.2.1- ¿Que es el Lenguaje C++?

El C++ es un lenguaje de programación, diseñado a mediados de los años 1980, por Bjarne Stroustrup, como extensión del lenguaje de programación C (6). Se puede decir que C++ es un lenguaje que abarca tres paradigmas de la programación: la programación estructurada, la programación genérica y la programación orientada a objetos (7).

4.2.2- ¿Que importancia tiene en Greenstone la utilización del lenguaje C++?

C++ combina la programación tradicional en C con programación orientada a objetos. Posibilita al Sistema Greenstone utilizar Perl (*Practical Extraction and Report Language*) que es un lenguaje de programación creado por Larry Wall en 1987 y que toma características del C, del lenguaje interpretado shell (sh), AWK, sed, Lisp y, en un grado inferior, muchos otros lenguajes de programación (8). Otro dato interesante es que Greenstone utiliza considerablemente la Biblioteca Estándar de Plantillas (STL), una biblioteca C++ de amplia difusión concebida por Silicon Graphics (http://www.sgi.com), y que es el resultado de numerosos años de investigación y desarrollo (5).

4.2.3- ¿Qué otros lenguajes de programación utiliza Greenstone?

4.2.3.1- SGML

El SGML (*Standard Generalized Markup Language*) es la especificación "padre" de todos los lenguajes de marcado. Es un estándar internacional (de la ISO) para la definición de métodos independientes de dispositivos y plataformas de representación de textos en forma electrónica (9). Es también una estructura que permite guardar información que facilita la manipulación, búsqueda precisa y el rearreglo de grandes repositorios de información. El SGML es una especificación general y presenta el concepto de tipo de documento. Así, cada tipo de documento tiene una especificación de estructura o DTD (*Documento Type Definition*). El SGML puede especificar "desde transcripciones de antiguos rollos sumerios hasta la documentación técnica de bombarderos Stealth y desde historias clínicas de pacientes hasta notas musicales", y por supuesto, incluye al hipertexto (10).

4.2.3.2- HTML

HTML (*HyperText Markup Language* o Lenguaje de Marcas Hipertextuales),es el lenguaje con el que se definen las páginas Web. Básicamente se trata de un conjunto de etiquetas que sirven para definir el texto y otros elementos que compondrán una página Web. El HTML es un lenguaje de marcación de elementos para la creación de documentos hipertexto (10). El archivo donde está contenido el código HTML es un archivo de texto, con una peculiaridad, que tiene extensión .html o .htm (es indiferente cuál utilizar). HTML es una aplicación de SGML conforme al estándar

internacional ISO 8879 (11).

4.2.3.3- XML

El lenguaje extensible de marcas, abreviado XML, describe una clase de objetos de datos llamados documentos XML y parcialmente describe el comportamiento de programas de computador que pueden procesarlos (11). XML es un perfil de aplicación o forma restringida de SGML (Standard Generalized Markup Language) [ISO 8879]. Por construcción, todo documento conforme con XML es conforme con SGML (11).

4.2.3.4- Lenguaje PHP

PHP es un lenguaje de programación muy potente que, junto con HTML, permite crear sitios Web dinámicos. PHP se instala en el servidor y funciona con versiones de Apache, Microsoft IIs, Netscape Enterprise Server y otros (12). La forma de usar PHP es insertando código PJHP dentro del código HTML de un sitio Web. Cuando un cliente (cualquier persona en la Web) visita la página Web que contiene éste código, el servidor lo ejecuta y el cliente sólo recibe el resultado. Su ejecución, es por tanto en el servidor, a diferencia de otros lenguajes de programación que se ejecutan en el navegador (13).

Por ultimo queremos referirnos en este apartado al Lenguaje de Macros

4.2.3.5- Lenguaje de Macros

El lenguaje de macros, se utiliza para conferir al sistema de Biblioteca Digital Greenstone un estilo coherente y crear interfaces en diferentes lenguas. La interacción con la biblioteca genera la estructura básica de las páginas Web y las macros del directorio GSDLHOME/macros aportan el contenido (mas adelante hablaremos del directorio GSDLHOME) (13).

Características Generales de Greenstone que merecen ser destacadas (14):

- Accesible vía Web por un navegador
- Servidores que funcionan en Windows, Unix o Mac
- Las colecciones pueden ser publicadas en CD-ROM/DVD
- De fácil instalación
- Interfaz del bibliotecario para crear y publicar las colecciones
- Búsquedas a texto completo y por campos
- Facilidades de búsqueda
- Basado en Metadatos (Dublin Core los recomendados)
- Acceso automático a todas las estructuras
- Utilización de Pluggins y formatos de Metadatos.
- Clasificadores para rastrear los documentos.
- Interfaces en varios idiomas. Visualización de documentos en varios idiomas.
- Utilización de Multimedia: imagen, video, audio en las colecciones.
- Uso de Multiformatos en los documentos y Metadatos.

Hemos mencionado entre estas características los conceptos de Pluggins y Metadatos, vamos a referirnos brevemente a ellos, pues mas adelante ampliaremos cuando hagamos referencia en el Capitulo de Creación de Colecciones

Un plug-in es un módulo de hardware o software que añade una característica o un servicio

específico a un sistema más grande (15). La idea es que el nuevo componente se enchufa simplemente al sistema existente. Por ejemplo, hay un gran número de plug-ins para el navegador Firefox que te permiten utilizar diversas herramientas y el plug-in de Flash, permite ver animaciones en flash en cualquier navegador (15).

El término metadato, fue acuñado por Jack Myers en la década de los 60 para describir conjuntos de datos. La primera acepción que se le dió (y actualmente la más extendida) fue la de dato sobre el dato, ya que proporcionaban la información mínima necesaria para identificar un recurso (16).

El sistema Greenstone en sus versiones mas recientes, ocupa unos 60 Mb de espacio de memoria en el disco duro. no es necesario que Ud. Instale el programa en el Disco C:\ de su equipo, puede hacerlo en el Disco D:\

4.2.3.6- Programas asociados al uso del software Greenstone

Sobre la plataforma Windows se requiere:

- Apache,
- Java RunTime,
- ImageMagick y
- Ghostscript.
- Otros programas: motor de búsqueda GM y GDBM.

4.2.4- Información Mínima sobre el Software Greenstone

4.2.4.1- Interoperabilidad: uso de estándares internacionales

¿A que nos referimos cuando utilizamos el termino "Estándar"?. Un Estándar se define como:

> *"... el común y repetido uso de reglas, condiciones, lineamientos o características de los productos o procesos relacionados y producción de métodos, y manejos relacionados en la práctica sistemática"* (17)

Relacionados con el software Greenstone podemos mencionar varios estándares y lineamientos que son importantes para así poder garantizar un valor presente y futuro de los documentos gestionados y su información, su expansión, integración e interconexión con o migración a otros sistemas. Estos estándares incluyen el Protocolo para Cosecha de Metadata (*Protocol for Metadata Harvesting: OAI-PMH*) de la *Open Archives Initiative*, así como OAIS, el modelo de referencia ISO para *Open Archival Information System*, definido por recomendación del Comité Consultivo para Sistemas de Datos Espaciales (CCSDS) y el Dublin Core para definición y estructura de metadatos, entre otros (20).

4.2.4.2- Servidor propietario del Greenstone

Greenstone incorpora un servidor que puede servir para cualquier colección basado en el Protocolo *Open Archives Protocol* para Metadata Harvesting (OAI-PMH) y hablaremos extensamente de este protocolo mas adelante.

Cualquier colección puede ser exportada a METS (*Metadata Encoding and Transmission Standard* en el *Greenstone METS Profile*, aprobado por METS Editorial Board y publicado en http://www.loc.gov/standards/mets/mets-profiles.html), Greenstone también puede asimilar documentos en formato METS .

El esquema METS es un estándar para codificar una descripción , administrativa , y estructural de los metadatos , con relación a los objetos de una biblioteca digital, expresados con el lenguaje XML y el Consorcio 3W (20).

El estándar es mantenido por *Network Development and MARK Standards Office* de la Biblioteca del Congreso ,y ha sido desarrollado como una iniciativa de la *Digital Library Federation. Metadata Encoding and Transmission Standard* (METS) *Official Web Site* http://www.loc.gov/standards/mets/mets-home.html

Podemos adelantar desde ahora la importancia de estos metadatos, los que son definidos por los usuarios interactivamente con la Interfaz del Bibliotecario. Estos sets de Metadatos están predefinidos:

- Dublin Core (calificados y no calificados)
- RFC 1807
- NZGLS (*New Zealand Government Locator Service*)
- AGLS (*Australian Government Locator Service*)

Ud. puede encontrar una definición precisa de cada uno de estos elementos en el Capitulo 18 Términos Conceptuales y Referenciales.

4.2.4.3- Plugins en Greenstone

Los plugins en Greenstone se utilizan para incorporar los documentos a las colecciones. Ampliaremos sobre esta idea cuando tratemos el Capitulo de Creación de Colecciones.

4.2.4.4- Función de administración

Una función "de administración" que establece los privilegios de cada tipo de usuario y de la cual hablaremos en el Capitulo de instalación del Software Greenstone.

Se pueden actualizar las colecciones y proponer nuevas colecciones en todo momento, sin apagar el sistema. Se establece mediante un proceso de la interfaz de usuario tomará nota (mediante sondeos periódicos) de las nuevas colecciones que aparezcan y las añadirá a la lista presentada al usuario. Esto será un aspecto a tratar en el Capitulo de Creación de Colecciones.

4.2.5- Programas asociados al uso del software Greenstone en Windows.

4.2.5.1- Ambiente de Ejecución Java

Para poder usar la Interfaz de Bibliotecario de Greenstone es necesario contar con una versión adecuada del Ambiente de Ejecución Java (JRE por sus siglas en inglés) (21).

4.2.5.2- Servidor Web Apache

Podemos trabajar con Greenstone de modo local o bien incorporando un servidor que nos permitirá trabajar en forma remota.

Las versiones mas recientes de Greenstone (a partir de la 2.8) cuentan con un servidor Web incorporado en el paquete de instalación. Pero necesitamos además un servidor Web externo.

El servidor Web Apache de http://www.apache.org.es libre y fácil de instalar (21).se trata de un servidor *open source* para los modernos sistemas operativos que incluye UNIX y Windows NT. El objetivo es proveer un servidor http, seguro, eficiente, y extensible con los estándares http vigentes

(21).

4.2.5.3- ImagemagicK

Es una robusta colección de herramientas y bibliotecas Open Source cuya licencia permite consultar, escribir y manipular una o varias imagenes (en 88 grandes formatos) incluyendo los formatos TIFF, JPEG, PNG, PDF, Photo CD, y GIF. ImageMagick es una aplicacion portable y compila bajo casi todos los sistemas que corren sobre 32-bit o 64-bit CPUs. ImageMagick se encuentra disponible para Unix o sistemas semejantes incluyendo Linux. Tambien corre sobre Winsows 95,98, ME, NT 4.0, 2000, and XP), Macintosh (5 y 9), VMS, and OS/2 (22).

4.2.5.4- Ghostscript

Ghostscript (escrito por Peter Deutsch, fundador de Aladdin Enterprises), es el programa intérprete por excelencia de documentos en formato PS (y también PDF). Lenguaje Postscript que fue desarrollado por la empresa Adobe, como formato de representación gráfica. La idea consiste en que una imágen, compuesta por figuras geométricas y por texto en tipos de letras y tamaños arbitrario puede ser representada de forma descriptiva. La formalización de esta descripción forma un lenguaje de programación (23).

Otros programas importantes que utiliza Greenstone: motor de búsqueda GM y GDBM.

4.2.5.5- MG (Managing Gigabytes)

Greenstone utiliza MG (Managing Gigabytes) para indizar y recuperar documentos, y su código fuente se incluye en el directorio GSDLHOME/packages. (Debemos precisar que GSDLHOME es el directorio principal del programa)

MG emplea técnicas de compresión para optimizar la utilización del espacio de disco sin alterar la rapidez de ejecución. En el caso de una colección de documentos en inglés, el texto comprimido y los índices suelen ocupar un tercio del espacio que necesitaría el texto original sin comprimir. La búsqueda y recuperación suelen ser más rápidas que las operaciones equivalentes en la versión sin comprimir porque requieren menos operaciones en el disco duro (5).

4.2.5.6- GDBM

Todas las versiones de Greenstone utilizan el Administrador de Base de Datos GNU, GDBM (GNU Database Manager). Viene con todas las versiones Windows de Greenstone y se instala automáticamente durante el procedimiento de instalación (20).

4.3- Greenstone es multiplataforma (Windows, Linux, Unix)

Greenstone funciona con los sistemas Windows (3.1/3.11, 95/98/Me, NT/2000) y Unix (Linux y SunOS). Cualquiera de estos sistemas puede utilizarse como servidor Web. Las colecciones no pueden crearse con versiones antiguas de Windows (3.1/3.11), pero pueden transferirse a esos sistemas colecciones previamente creadas (24).

Greenstone corre en todas las versiones de Windows, Unix/Linux, y Mac OS-X. Es de muy fácil instalación. Para la instalación en Windows no es necesaria ninguna configuración especial, el usuario final puede instalar el software en su computadora personal. Los usuarios institucionales pueden instalar el software en un servidor Web y entonces interoperar con los estándares Web del software servidor (que como hemos visto previamente se recomienda sea el apache). (25).



Figura 4.1. Interfaz de usuario en Greenstone. Visualización colección demo

Esta es la página inicial de Greenstone y es la interfaz denominada del lector o usuario, en su producto original, pero la misma puede ser modificada de acuerdo a las necesidades de los responsables del proyecto de implementación de la biblioteca de que se trate.

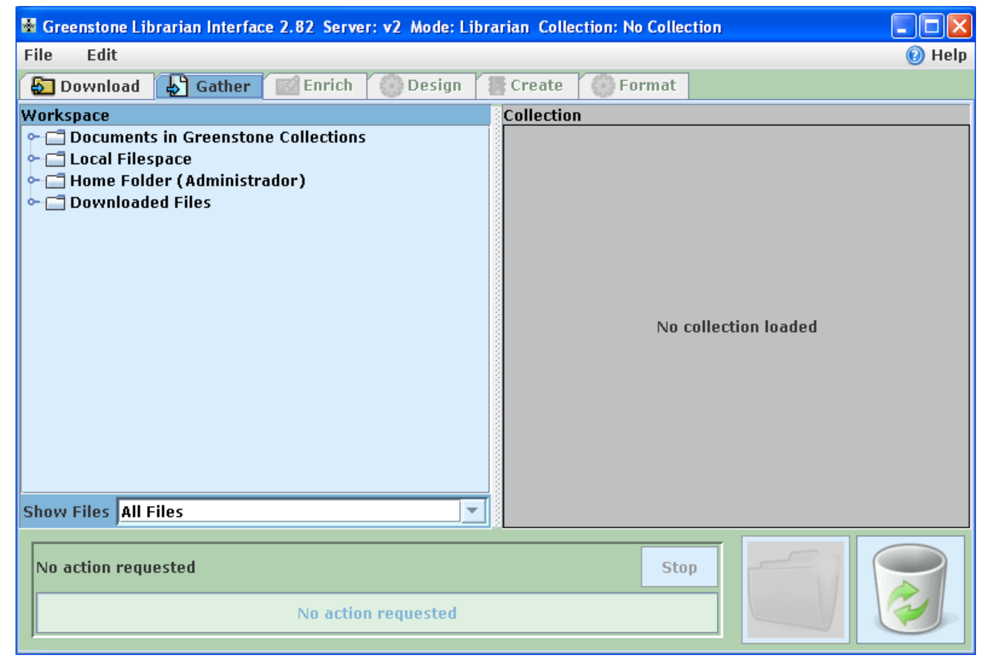


Figura 4.2. Interfaz gráfica de trabajo del bibliotecario (GLI)

La interfaz del bibliotecario de Greenstone (GLI por sus siglas en inglés) es una herramienta gráfica (como hemos explicado previamente) que le ayuda a construir las colecciones de las bibliotecas digitales dándole acceso a todas las funciones de Greenstone por medio de una interfaz fácil de usar con la que sólo tiene que apuntar y hacer clic.

Situados los antecedentes básicos necesarios, en el próximo capitulo describirnos las particularidades de instalación del software.

4.4- Consideraciones Finales

El software Greenstone califica como un sistema en el que interactúan varias aplicaciones,

lenguajes de programación y herramientas como hemos destacado previamente que hacen posible codificar de modo único e irrepetible toda la información que se introduce al sistema lo que posibilita un nivel de recuperación exhaustivo y pertinente en cualquiera de las modalidades que se utilice lo cual lo coloca en un nivel de efectividad extraordinario.

4.5- Referencias Bibliográficas

(1) Greenstone .org. Biblioteca Digital Greenstone. Guía del Programador .Configuración del sitio Greenstone. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm>

(2) Tramullas J. Biblioteca Digital Greenstone. Capitulo 10. En.Software Libre para Servicios de Información [Libro en Internet] Digital. Madrid: Pearson Prentiss Hall, 2006. Disponible en: <http://tramullas.com>

(3) Programación en Castellano, S.L. .Curso práctico de Corba en GNU/Linux. [Sitio Web en Internet]. 2008. Disponible en: <http://www.programacion.com/templates/default/css/estilo.css>

(4) osmosislatina.com. RMI y CORBA: Una base para aplicaciones Distribuidas. [Sitio Web en Internet] 2005/09/07. Disponible en: <http://www.osmosislatina.com/css/osmosis2006.css>

(5) Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 2 Exportación de mis documentos. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/style.css>

(6) PergaminoVirtual.Glosario [Sitio Web en Internet] 2007-06-12 , Disponible en: <http://www.pergaminovirtual.com.ar/definicion/default.css>

(7) Greenstone Wiki. Biblioteca Digital Greenstone. Guía de Usuario. Apéndice B: Glosario de Términos [Sitio Web en Internet], Marzo 2006. Disponible en: <http://greenstonewiki.cs.waikato.ac.nz/wiki/gsdoc/GNUFDL.html>

(8) PergaminoVirtual Definición de PERL [Sitio Web en Internet] 2007-06-12 <http://www.pergaminovirtual.com.ar/definicion/default.css>

(9) Gaiasur.com. Los lenguajes de marcado [Sitio Web en Internet]. 2008. Disponible en:< <http://fractal.gaiasur.com.ar/infoteca/seminarios/pub-/conceptos/hipertexto.html>

(10) Álvarez M.A. Llegó el momento de hablar sobre HTML, el lenguaje con el que crean las páginas Web. [Sitio Web en Internet]. 29/9/01. Disponible en: <http://www.desarrolloweb.com>

(11) W3C El lenguaje extensible de marcas (XML) 1.0 - Extensible Markup Language (XML) Recomendación de la W3C [Website in Internet], - Febrero 1998. Editores: Tim Bray (Textuality y Netscape) Jean Paoli (Microsoft) C. M. Sperberg- McQueen (University of Illinois at Chicago) . Traducido al español por: Fabio Arciniegas A. 1999/08/17. Disponible en: <http://www.sidar.org>

(12) masadelante.com. ¿Que significa PHP? - Definición y explicación del término php. [Sitio Web en Internet] 1999. Disponible en: <http://www.masadelante.com/favicon.ico>

(13) Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3. El Sistema de Ejecución de Greenstone [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/style.css>

(14) Witten I. Introducing the Greenstone Digital Library Software [Website in Internet], 2009.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

Computer Science Department- Waikato University New Zealand. Disponible en: <http://www.nmis.cnv.it/casarose/PUB.DL/greenstone/greenstone_overviw_slides.pdf>

(15) Masadelante.com - ¿Que es un plug-in? - Definición de plug-in [Sitio Web en Internet] 1999.Disponible en: <http://www.masadelante.com/faqs/plug-in>

(16) Senso J. A. de la Rosa A. El concepto de metadato. Algo más que descripción de recursos electrónicos [Sitio Web en Internet] 2003 <http://www.scielo.br/pdf/ci/v32n2/17038.pdf>

(17) Library of Congress. Technical Reports and Standards [Website in Internet]. 2005 Available from: <http://www.loc.gov/global/stdLauncher.js>

(18) Miranda C. Sistemas de Gestión de Documentos bajo Licencia de Código Libre [Sitio Web en Internet] 08 Junio 2009. Disponible en:<http://opengraphprotocol.org/schema/xmlns:fb>

(19) Library of Congress. Metadata Encoding and Transmission Standard (METS) Official Web Site [Website in Internet], November 25, 2009. Available from: <http://www.loc.gov/standards/mets/mets-home.html>

(20) Greenstone.org. Guía de Instalación Greenstone. Apendice: Programas Asociados.[Sitio Web en Internet]. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone>

(21) Apache. Http Server Project. The number one HTTP server on the Internet.[Website in Internet] 2010/04/27. Avilable from: <http://httpd.apache.org>

(22) Imagemagick.org.imagemagick. [Website in Internet] 2002. Available from: <http://www.imagemagick.org>

(23) Indaganda.com - Definición de Ghostscript y Información sobre Ghostscripthttp [Sitio Web en Internet] 2001. Disponible en: <http://www.indaganda.com/definir/Ghostscript>

(24) Biblioteca Digital Greenstone. Guía de Usuario Características del Programa [Sitio Web en Internet] Marzo 2006. Disponible en: <http://greenstonewiki.cs.waikato.ac.nz/wiki/gsdoc/GNUFDL.html>

(25) Greenstone.org Factsheet :Greenstone Digital Library Softwar [Website in Internet] 11/2000. Available from:<<http://www.greenstone.org/factsheet>

CAPÍTULO 5

5- Particularidades en la instalación del software Greenstone

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH)

Ms C Ricardo Fernández Rodríguez (ricardo@cecam.sld.cu) Profesor Auxiliar del I.S.C.M.H. Máster en Informática en Salud. Centro de Cibernética Aplicada a la Medicina (CECAM) I.S.C.M.H. Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba. CP: 11600, Telf.: (537) 271-1354

Palabras clave: Desarrollo Sostenible, Economía Digital, Estrategias de Inclusión Digital. , Sociedad de la Información, Software Libre, UNESCO.

5.1- Introducción

Geenstone es un software open source (o software de código abierto) que se distribuye de acuerdo a los términos de la Licencia Publica General (GNU) http://www.gnu.org/gnu/gnu-history .es.html

Existen sendas versiones de Greenstone disponibles para Windows y Unix, como binarios y en forma de código fuente (1).

5.2- Requerimientos de hardware previos a la instalación

Es importante considerar que, al tratarse de un servicio que se ejecuta a petición del servidor Web, la instalación debe hacerse desde una cuenta con derechos para modificar determinados ficheros de configuración y permisos de ejecución (2).

En el caso de instalar sobre Linux, la instalación deberá hacerse desde una cuenta de administrador o root; para el caso de Windows debe hacerse desde el administrador o un usuario del grupo de administradores. Aunque puede instalarse sin ser administrador, es conveniente hacerlo correctamente desde un principio para evitar inconvenientes posteriores con los permisos.

Es importante recordar que en informática (3), un servidor es un tipo de software que realiza ciertas tareas en nombre de los usuarios. El término servidor ahora también se utiliza para referirse al ordenador físico en el cual funciona ese software, una máquina cuyo propósito es proveer datos de modo que otras máquinas puedan utilizar esos datos (para mayor ampliación sobre el término consulte el Capitulo 18 Términos Conceptuales y Referenciales).

El programa Greenstone funciona en diferentes plataformas y con distintas configuraciones (4). Existen sendas versiones de Greenstone disponibles para Windows y Unix, como binarios y en forma de código fuente. La interfaz de usuario de Greenstone utiliza cualquiera de los dos navegadores Web siguientes: Netscape Navigator o Internet Explorer (versión 4.0 o superior en ambos casos).

Cuando usted se enfrenta al instalador de este programa cabe destacar que se visualizaran los objetos de forma diferente si quiere utilizar una versión para Windows o para Linux. Si en su computadora no esta instalado el Sistema Operativo Linux y solo cuenta con el Sistema Operativo Windows, los verá así:



Figura 5.1 Vista de los logotipos de la versión Greenstone 2.x

Los iconos de la izquierda representan las aplicaciones del programa y los de la derecha representan las aplicaciones que dentro del núcleo del programa instalarán el servidor en uno u otro Sistema Operativo ya que el Greenstone a partir de la versión 2.82 que es sobre la que desarrollaremos este libro cuenta con un pequeño servidor Web apache que puede servir para el caso de una pequeña red, aunque para configurar varias maquinas deberá instalar en su maquina servidor el apache como programa, En nuestro caso utilizaremos la versión para Windows.

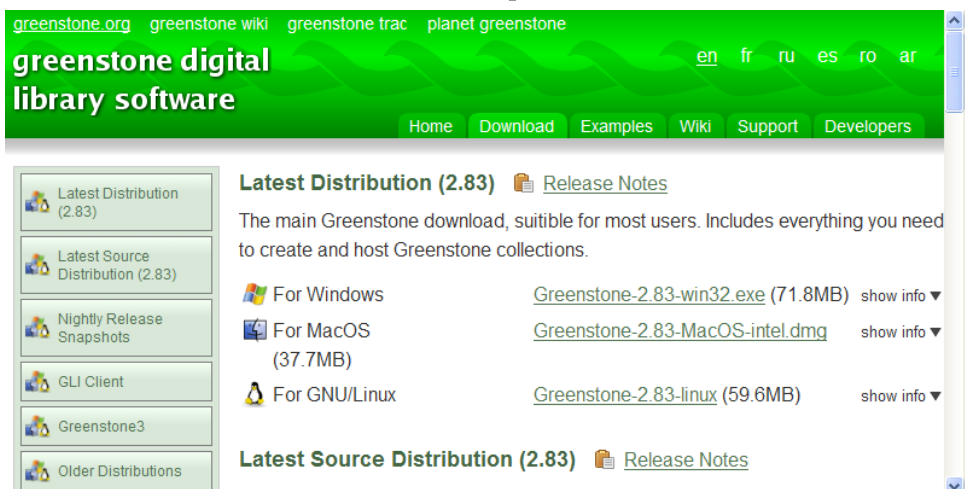


Figura 5.2 Página inicial para la descarga del Greenstone

¿Donde puede Ud. obtener la versión de Greenstone que desee instalar en su computadora? Este es el sitio desde donde se puede descargar el programa libremente: http://www.greenstone.org/blog

5.2.1- Instalación de Apache

La Fundación del Software Apache provee el soporte necesario para la utilización de esta

herramienta y constituye un grupo de proyectos open source, es decir de código abierto, que se caracteriza por la colaboración. Los procesos basados en consenso y el uso de una licencia pragmática para el software además del deseo de crear un software de calidad.

La Fundación provee la infraestructura para varios de los proyectos Apache.

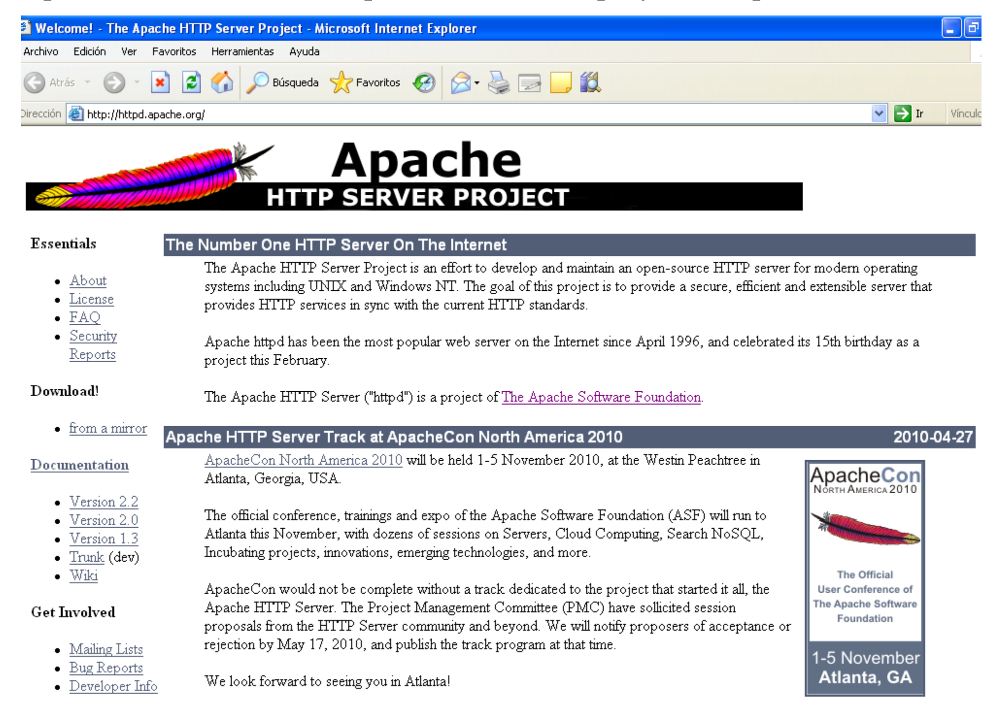


Figura 5.3. Página inicial del Apache HTTP server project

El proyecto servidor apache constituye un esfuerzo para desarrollar y mantener un servidor *open source* para los modernos sistemas operativos que incluye UNIX y Windows NT. El objetivo de este proyecto es proveer un servidor http, seguro, eficiente, y extensible con los estándares http vigentes.

Apache http ha sido el servidor Web más popular y es el que utilizaremos para soportar los servicios Web de la biblioteca de Greenstone.

Todos los servidores Web utilizan característicamente la URL especial “localhost” que representa la computadora en que está funcionando el servidor Web. Así, cuando se instala un servidor Web, se puede acceder a los documentos HTML tecleando la URL http://localhost en un navegador.

Si su computadora tiene instalado un nombre de dominio, éste se utiliza en lugar del local host para poder identificar su computadora desde sitios remotos (5).

Ud. puede obtener este programa desde varios sitios. En la Figura 5.4 se les muestran varias opciones.

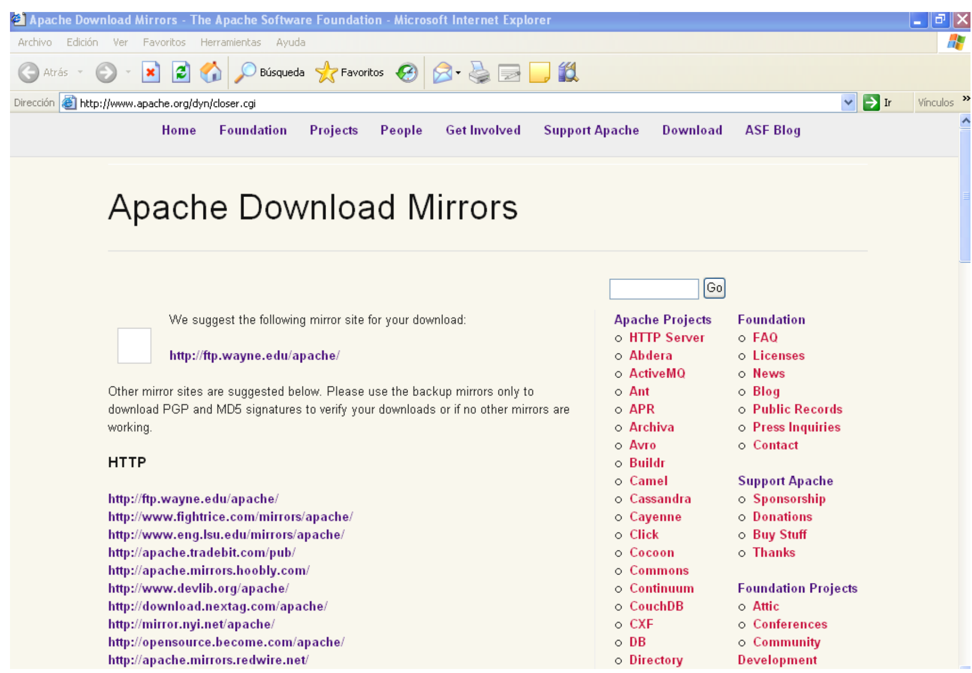


Figura 5.4. Espejo donde se puede descargar el Apache.

El servidor Web Apache suele instalarse en C:\Program Files\ApacheGroup\Apache y se configura de tal modo que el directorio cgi-bin esté en el subdirectorio \cgi-bin y la raíz de documentos es el subdirectorio \htdocs.

¿De que estamos hablando cuando decimos subdirectorio cgi-bin? cgi-bin es un directorio desde el cual el servidor Web trata documentos como programas ejecutables.

La directiva ScriptAlias de Apache se utiliza entonces para crear un directorio cgi-bin. Cabe señalar que esta directiva puede hacer de cualquier directorio un directorio CGI ejecutable (5).

Una vez instalado, Apache tiene un directorio cgi-bin en C:\Program Files\Apache Group\Apache\cgi-bin. Esto significa que si se le presenta la URL http://localhost/cgi-bin/hello, el servidor Web intentará ejecutar un archivo llamado hello en el directorio antes mencionado (5)

Tenga en cuenta Ud. que el directorio raíz de los documentos es la raíz de la estructura de directorios de su servidor Web.

Repetimos, una vez instalado, Apache dispone como raíz de documentos de C:\Program Files\Apache Group\Apache\htdocs. Esto significa que si se le presenta la URL http://localhost/hello.html, el servidor Web intentará recuperar un archivo llamado hello.html desde dentro del directorio antes mencionado (5)

Ampliaremos este tema en el capitulo que trata sobre la Configuración del servidor de Greenstone para usuarios en la segunda sección de este libro.

Una vez obtenida la aplicación esta será la característica del contenido de las carpetas que integran el programa apache

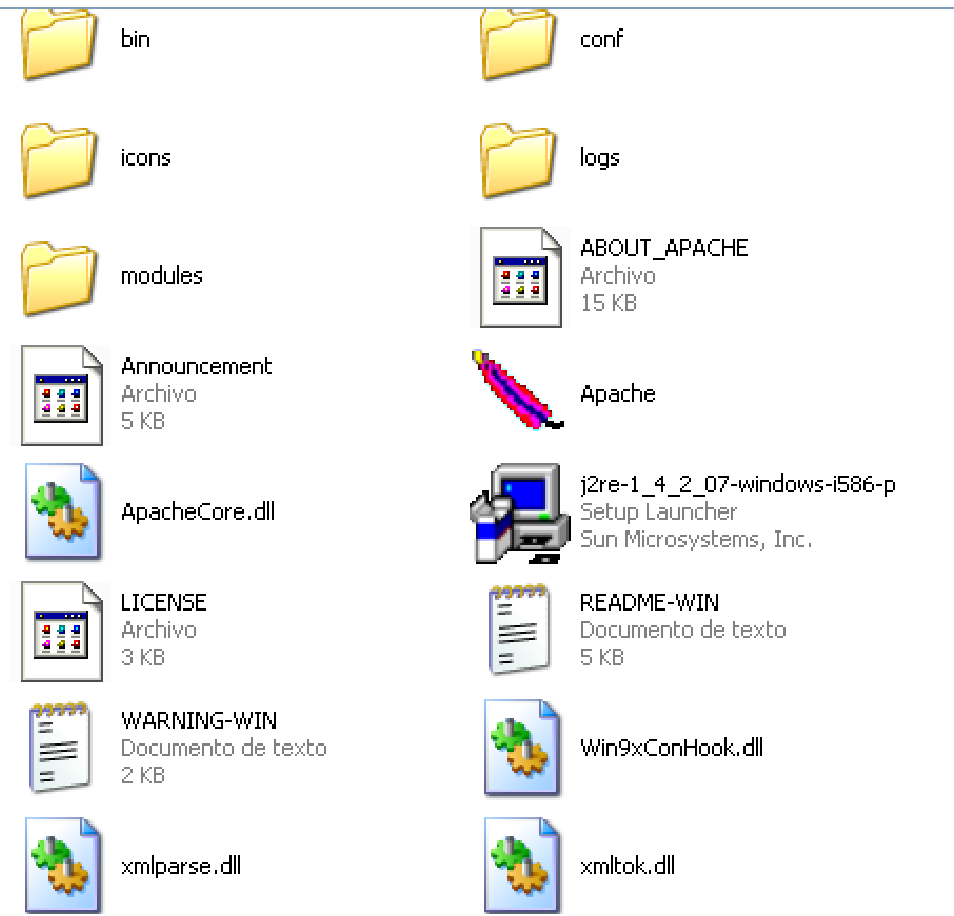


Figura 5.5. Contenido de las carpetas que integran el programa Apache

Cuando Ud. instala este software en un equipo este se convertirá en un servidor que podrá brindar determinadas prestaciones.

Debemos destacar que en la carpeta www se alojaran los sitios Web que Ud. manejará o sea esta carpeta representa la carpeta de publicación del Apache.

También debemos destacar que en la barra de tareas de su computadora aparecerá el símbolo de apache de color rojo o verde en dependencia de si su programa este en activo o no.



Queremos puntualizar que apache integra a su vez a la base de datos MySQL, y el lenguaje PHP de forma característica.

MySQL, es la base de datos open source más popular. Con una velocidad superior, confiabilidad y de facil uso que la ha convertido en el software de preferencia en la Web MySQL es un elemento importante de LAMP (Linux, Apache, MySQL, PHP / Perl / Python), la mas pujante empresa de software open source.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

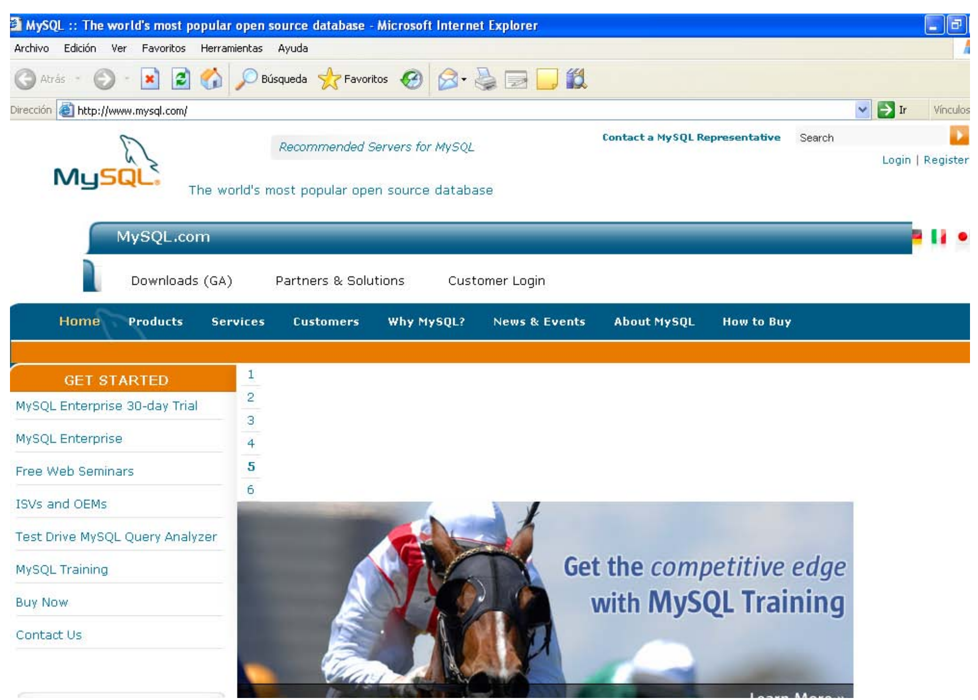
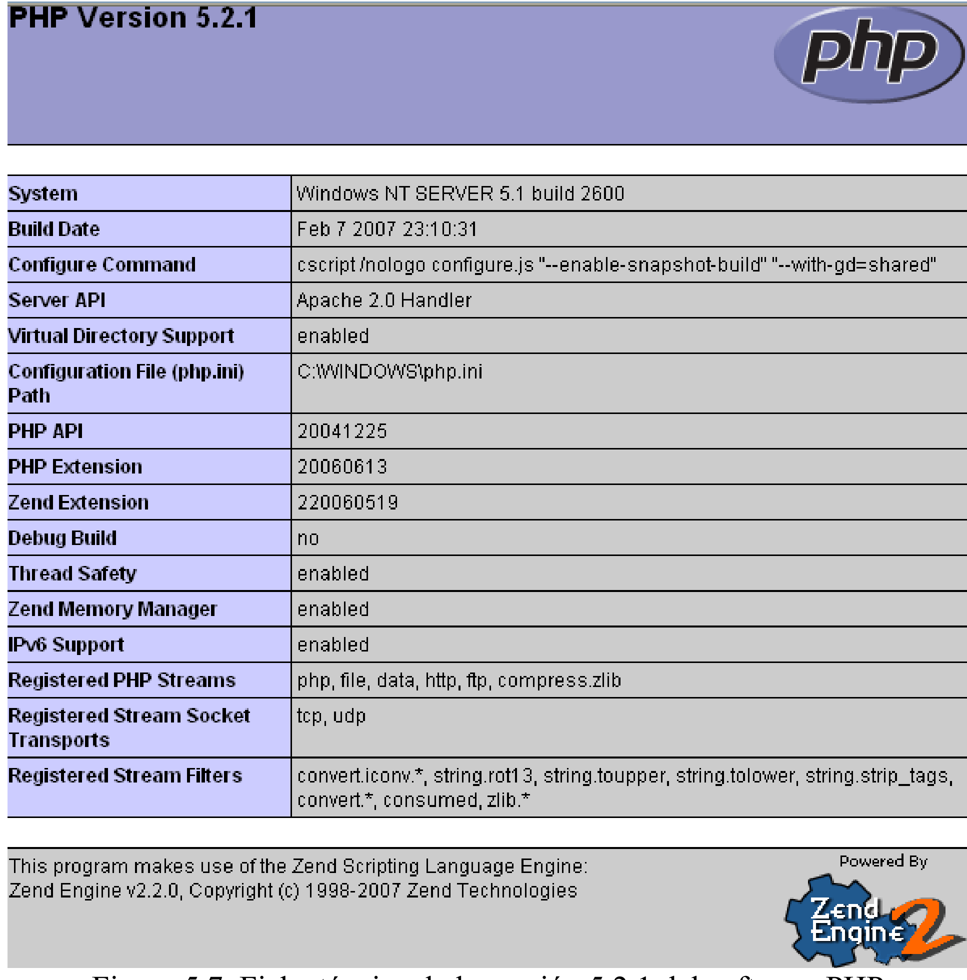


Figura 5.6. Interfaz inicial del sitio MySQL

Y aquí les mostramos como ejemplo una ficha técnica con algunas de las características de PHP, software libre, que Ud. puede redistribuir o modificar bajo los términos de la licencia PHP.


PHP Version 5.2.1

System	Windows NT SERVER 5.1 build 2600
Build Date	Feb 7 2007 23:10:31
Configure Command	cscript /nologo configure.js "--enable-snapshot-build" "--with-gd=shared"
Server API	Apache 2.0 Handler
Virtual Directory Support	enabled
Configuration File (php.ini) Path	C:\WINDOWS\php.ini
PHP API	20041225
PHP Extension	20060613
Zend Extension	220060519
Debug Build	no
Thread Safety	enabled
Zend Memory Manager	enabled
IPv6 Support	enabled
Registered PHP Streams	php, file, data, http, ftp, compress.zlib
Registered Stream Socket Transports	tcp, udp
Registered Stream Filters	convert.iconv.*, string.rot13, string.toupper, string.tolower, string.strip_tags, convert.*, consumed, zlib.*

This program makes use of the Zend Scripting Language Engine:
Zend Engine v2.2.0, Copyright (c) 1998-2007 Zend Technologies

Figura 5.7. Ficha técnica de la versión 5.2.1 del software PHP

5.2.2- Java RunTime

Para poder usar la Interfaz de Bibliotecario de Greenstone es necesario contar con una versión adecuada del Ambiente de Ejecución Java (JRE por sus siglas en inglés)

El software de Java para el equipo o el entorno de ejecución Java se denomina también ejecución Java, entorno de ejecución, ejecución, JRE, máquina virtual Java, máquina virtual, Java VM, JVM, VM o descarga de Java. La versatilidad y eficiencia de la tecnología Java, la portabilidad de su plataforma y la seguridad que aporta, la han convertido en la tecnología ideal para su aplicación a redes.

Desde este sitio puede Ud. obtener la aplicación que necesita:

http://www.java.com/es/download/manual.jsp#win

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

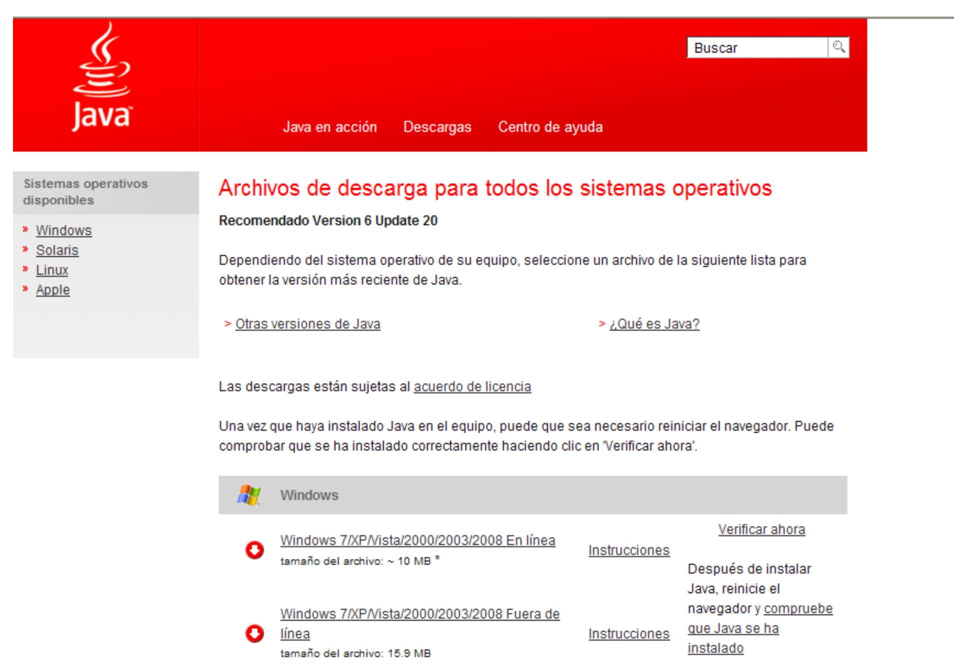


Figura 5.8. Interfaz del sitio Java de donde puede descargarse el software.

La plataforma Java es el nombre de un entorno o plataforma de computación originaria de Sun Microsystems capaz de ejecutar aplicaciones usando el lenguaje de programación java u otros lenguajes que compilen a bytecode y un conjunto de herramientas de desarrollo. En este caso, la plataforma no es un hardware específico o un sistma operativo, sino más bien una maquina virtual encargada de la ejecución de las aplicaciones, y un conjunto de bibliotecas estándar que ofrecen una funcionalidad común (6).

5.2.3- ImageMagick

Como ya habíamos mencionado en el capitulo anterior, ImageMagickTM 5.5.7 es una robusta colección de herramientas y bibliotecas cuya licencia permite consultar, escribir y manipular una o varias imágenes (en 88 grandes formatos) incluyendo los formatos TIFF, JPEG, PNG, PDF, PhotoCD, y GIF (http://www.imagemagick.org)

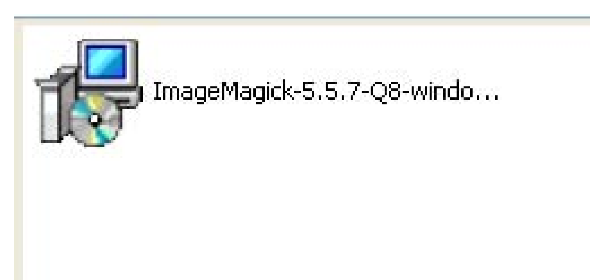


Figura 5.9. Apariencia de la aplicación del programa ImageMagick

Las operaciones del procesamiento de imágenes se encuentran disponibles desde la linea de comando , y también desde las interfaces de programación de C, C++, Perl, or Java. ImageMagick se encuentra disponible para Unix o sistemas semejantes incluyendo Linux. Tambien corre sobre Winsows 95,98, ME, NT 4.0, 2000, and XP), Macintosh (9 y 10), VMS, and OS/2.

Veamos en detalle el proceso de instalación que debe preceder a la instalación de Greenstone

Esta es primera interfaz de instalación del programa:

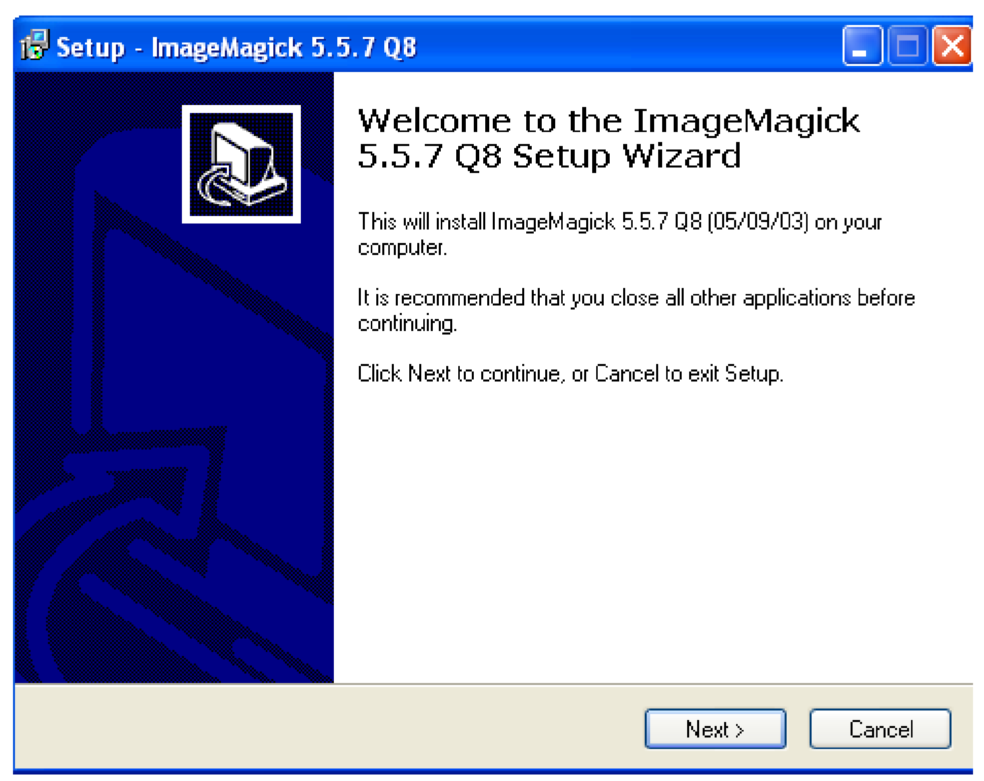


Figura 5.10. Primera interfaz de la aplicación del programa imageMagik.

La siguiente pantalla nos muestra las condiciones de uso del programa:

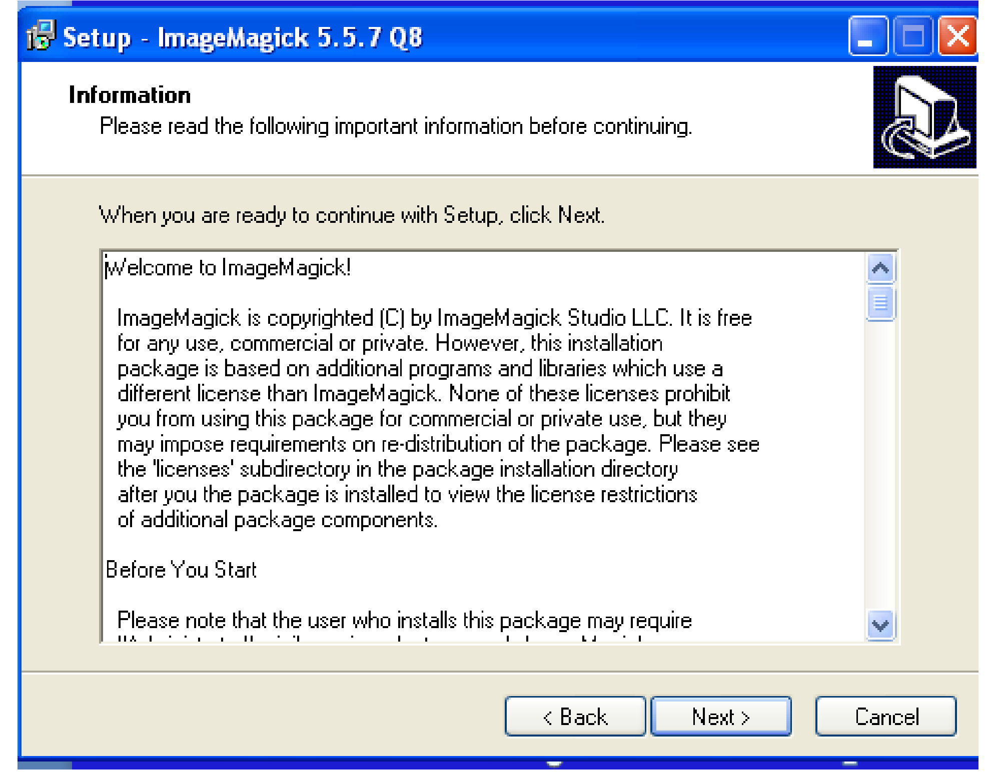


Figura 5.11. Información sobre las condiciones de uso del programa ImageMagick

A continuación se nos ofrece la posibilidad de seleccionar el lugar donde alojar el programa

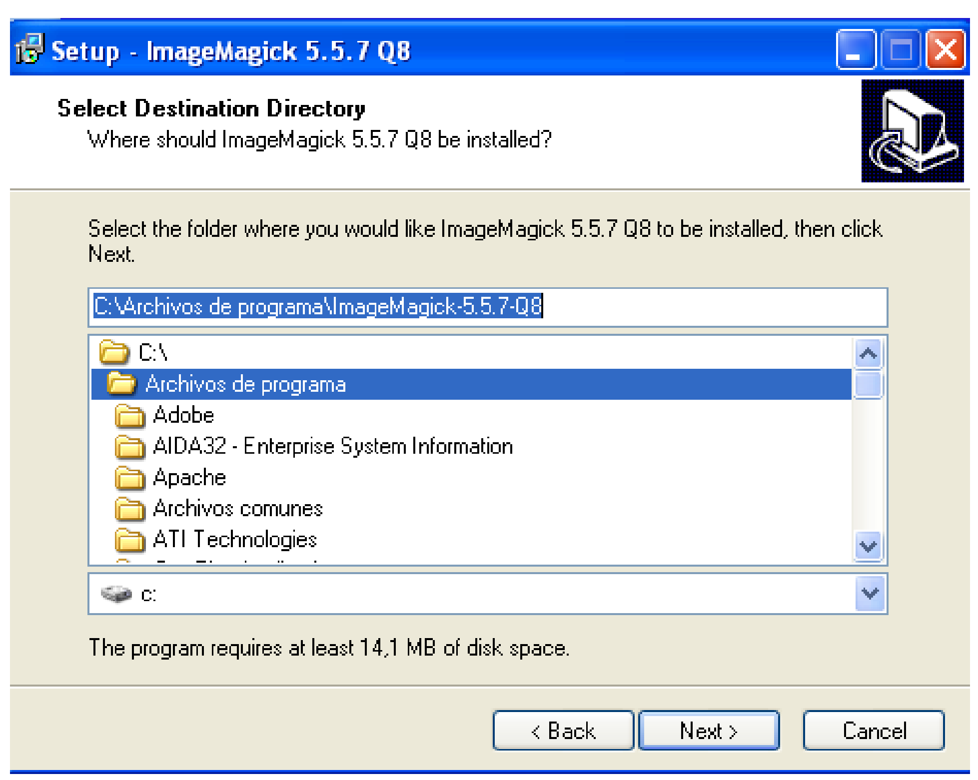


Figura 5.12. Opción de seleccionar el lugar donde hospedar el programa ImageMagick

Y la posibilidad de modificar las condiciones de trabajo del software.

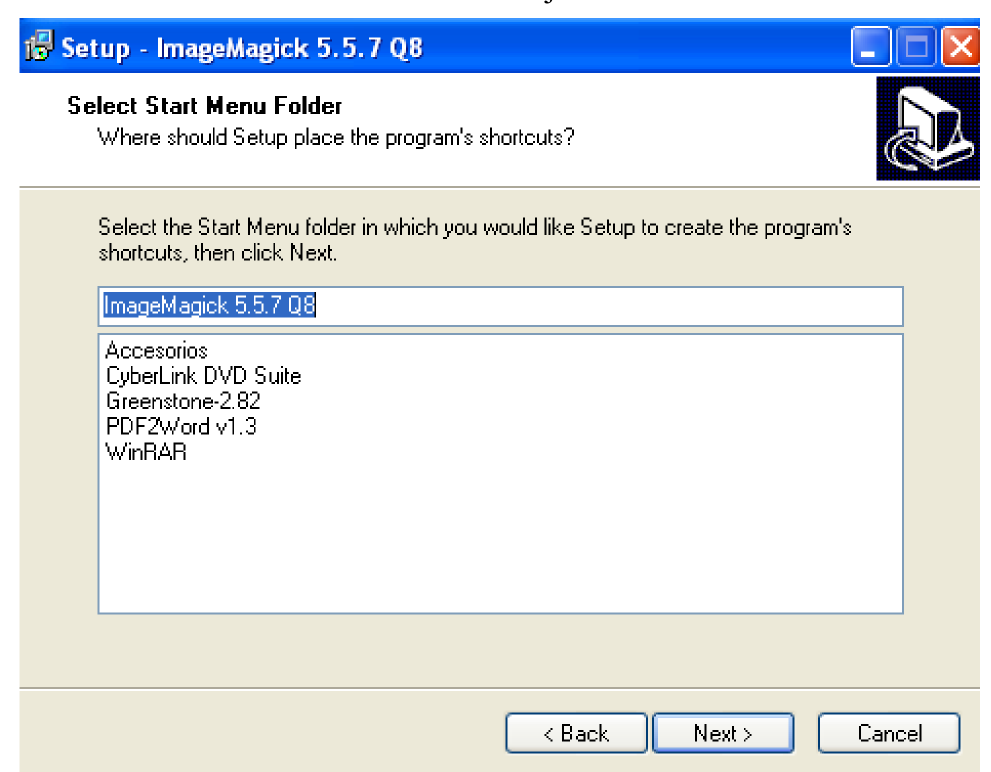


Figura 5.13. Opción de seleccionar las condiciones de trabajo del programa ImageMagick

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

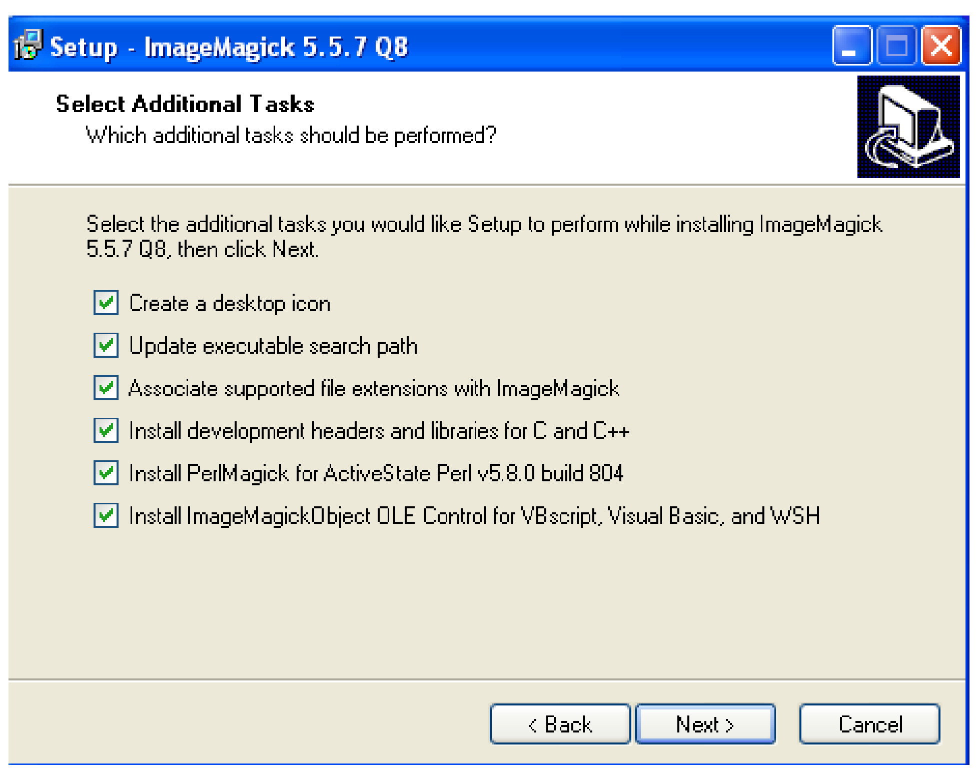


Figura 5.14. Opciones de seleccionar las tareas y accesorios del programa ImageMagick

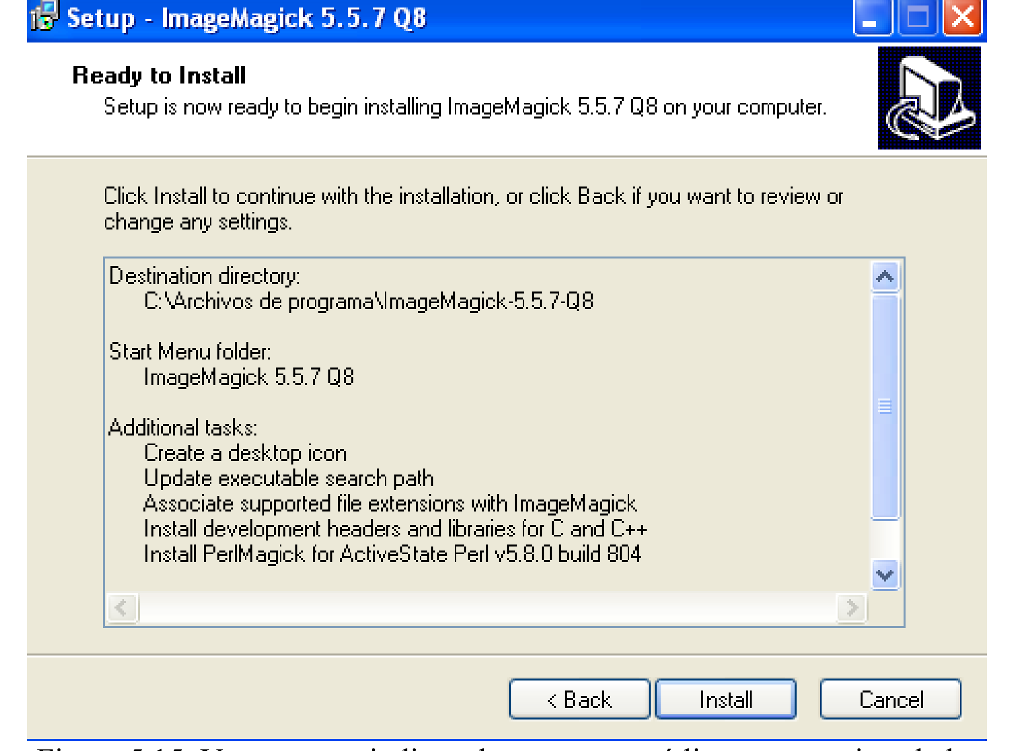


Figura 5.15. Ventana que indica: el programa está listo para ser instalado.

Es posible visualizar el procedimiento de instalación.

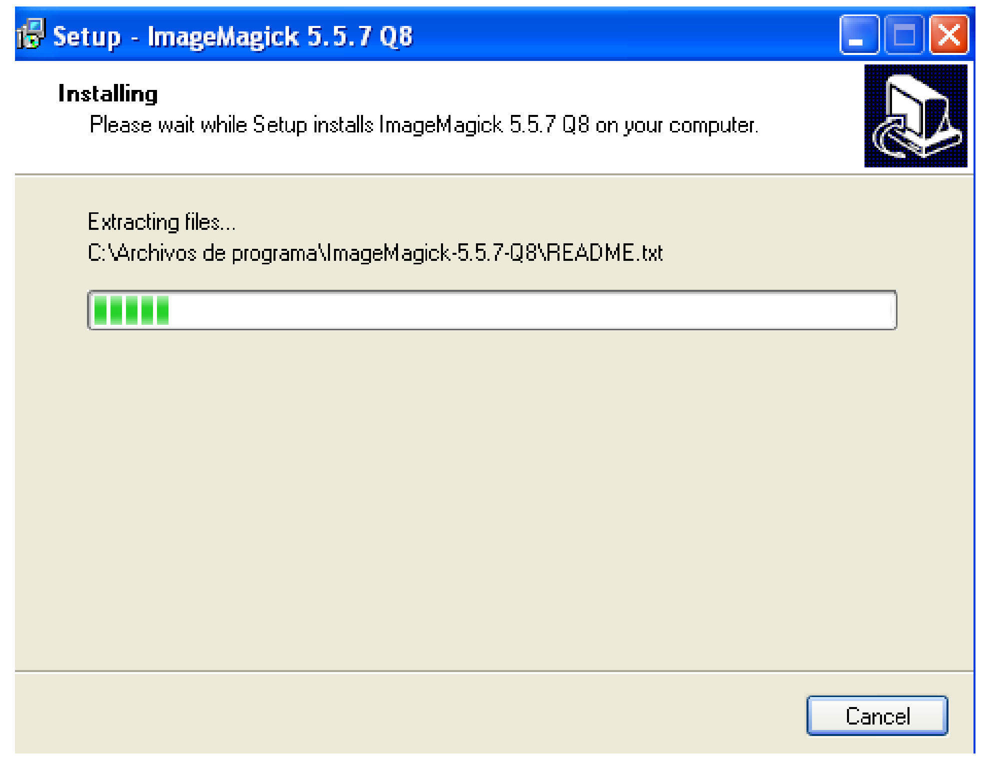


Figura 5.16. Opción de visualizar el proceso de instalación del programa ImageMagick

Una vez finalizada la instalación del programa recibirá un aviso como el siguiente:

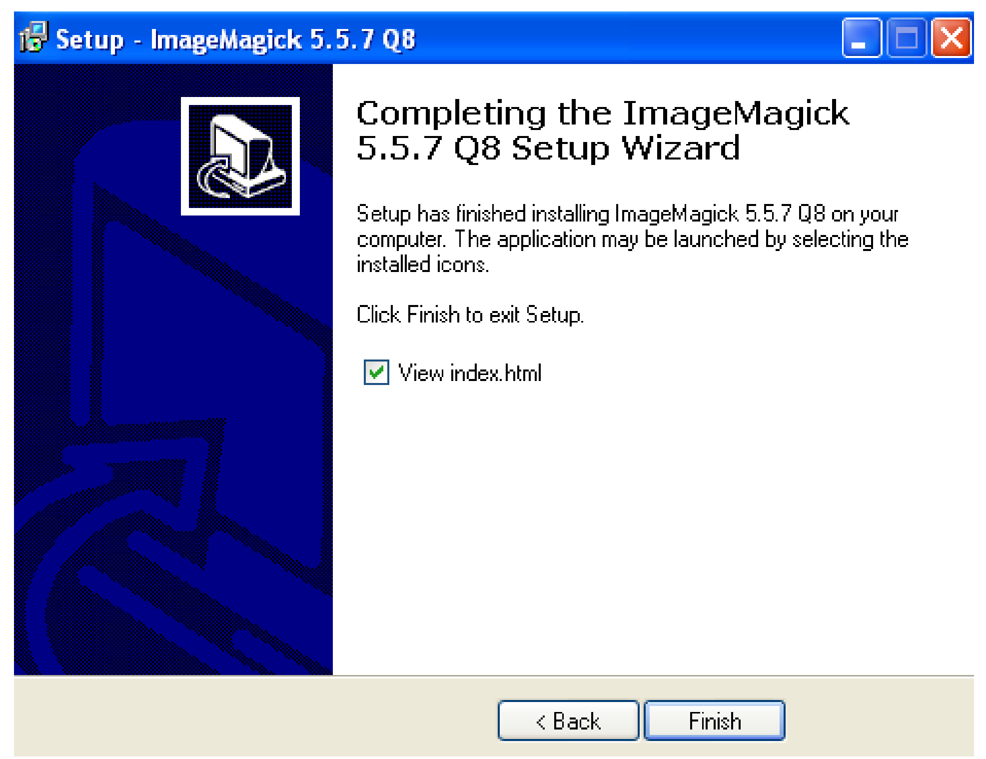


Figura 5.17. Visualización del proceso de instalación del programa ImageMagick y culminación del proceso.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

A continuación les mostramos la representación de las carpetas que contiene el programa:

Figura 5.18. Visualización del contenido de la carpeta del programa ImageMagick

Una vez instalados los programas previos podremos dirigir nuestro objetivo hacia la instalación de Greenstone.

5.3- Instalación del software Greenstone

Para proceder a la instalación más simple, basta con aceptar el parámetro por defecto en cada etapa, pulsando el botón Siguiente (Next si ha elegido la opción en inglés). Consideremos ahora el proceso real de instalación del programa.

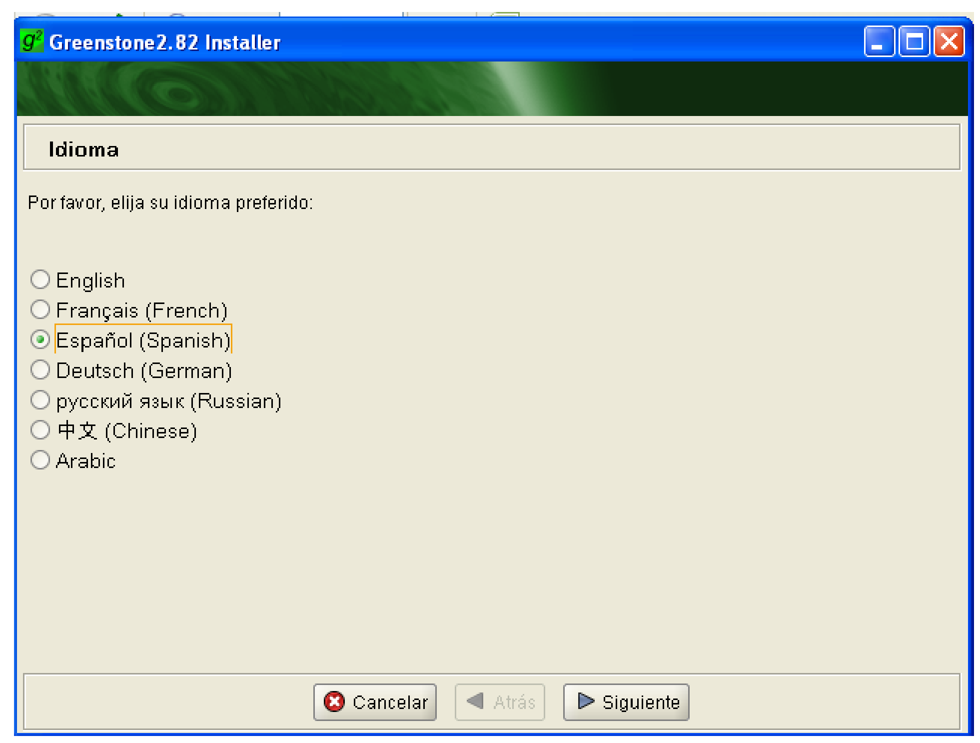


Figura 5.19. Pantalla que interroga al usuario sobre su preferencia de idioma de instalación

5.3.1- Idioma

Una vez usted haga doble clic sobre el icono del instalador aparecerá una primera pantalla como esta que le interrogará sobre su preferencia en cuanto al idioma con el cual quiere trabajar (Figura 5.19).

5.3.2- Información sobre programas

La siguiente pantalla que le mostrara una breve información sobre los programas que contiene la aplicación (Figura 5.20).

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

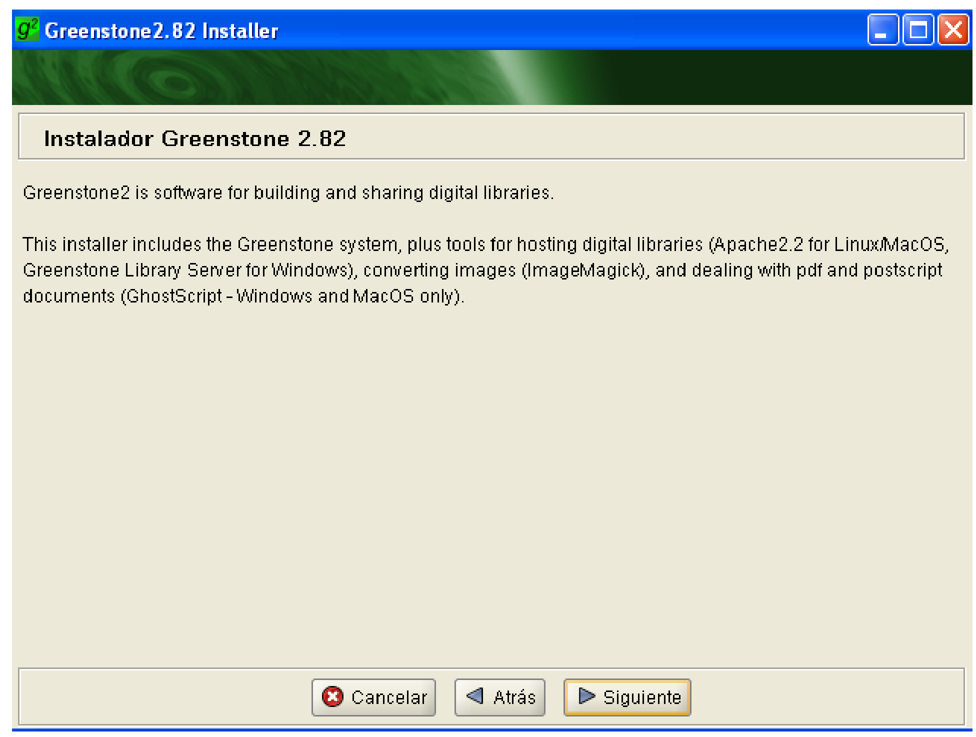


Figura 5.20. Pantalla informa sobre los programas que contiene la instalación.

5.3.3- Licencia

En la siguiente pantalla se registran las características del programa en cuanto a la licencia GNU (Figura 5.21).

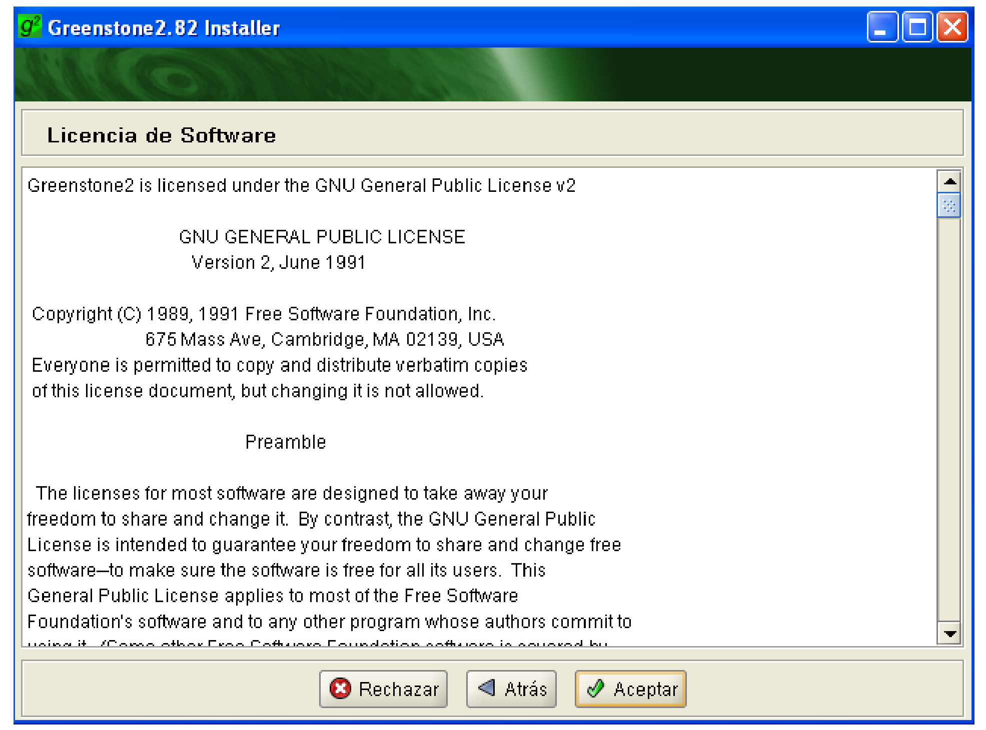


Figura 5.21. Características de la licencia GNU del programa Greenstone.

En la próxima pantalla (Figura 5.22) se hará visible el proceso de instalación y es aquí donde usted

debe definir donde ubicará el programa. El botón examinar en el top derecho le permitirá rastrear el contenido de su ordenador y seleccionar entonces la ubicación del programa. Es recomendable crear previamente una carpeta en el sitio en que haya decidido instalar el programa y proceder posteriormente a colocarlo allí.

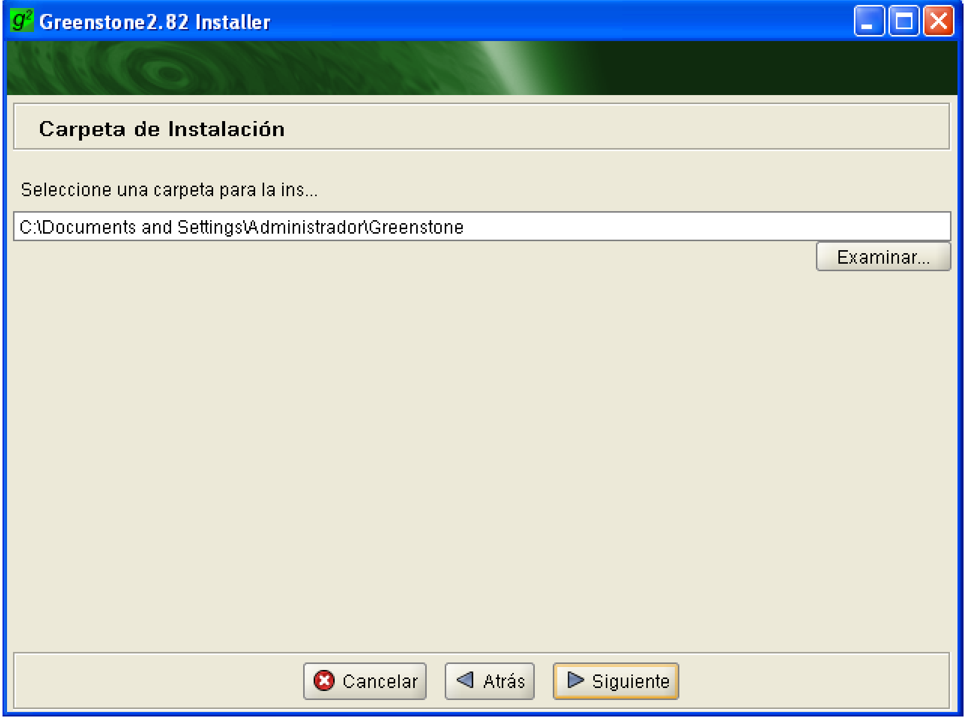


Figura 5.22. Opción de selección de la carpeta para la instalación del programa Greenstone.

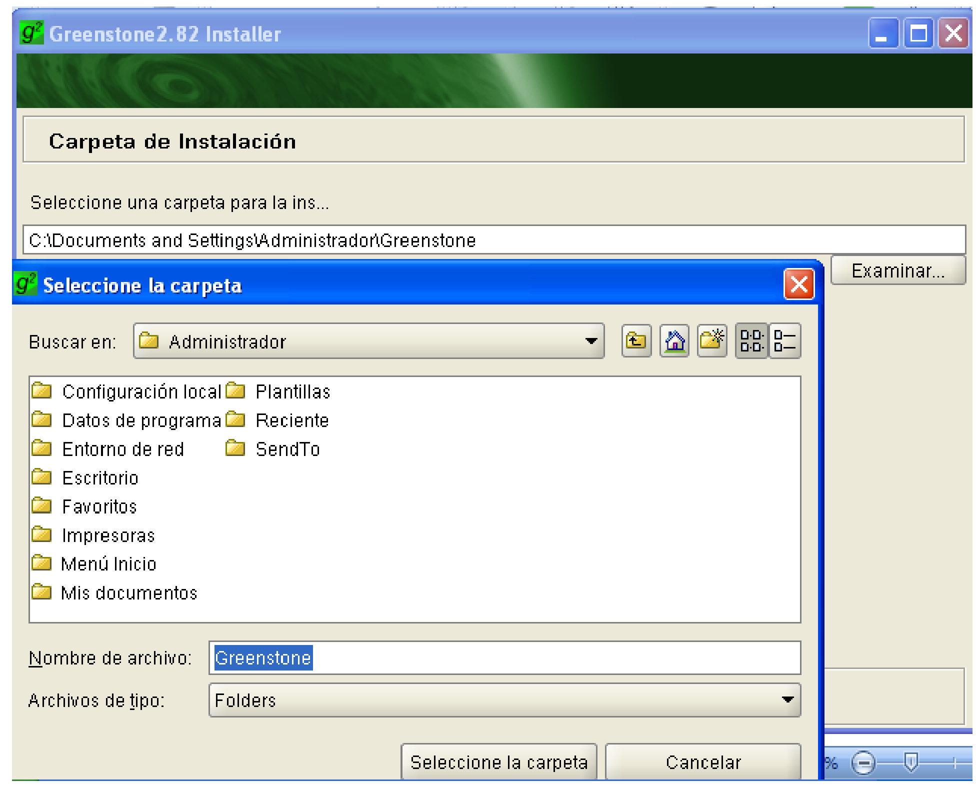


Figura 5.23. Creación previa de la carpeta donde se instalará el programa Greenstone.

En este caso se ha seleccionado el disco C.

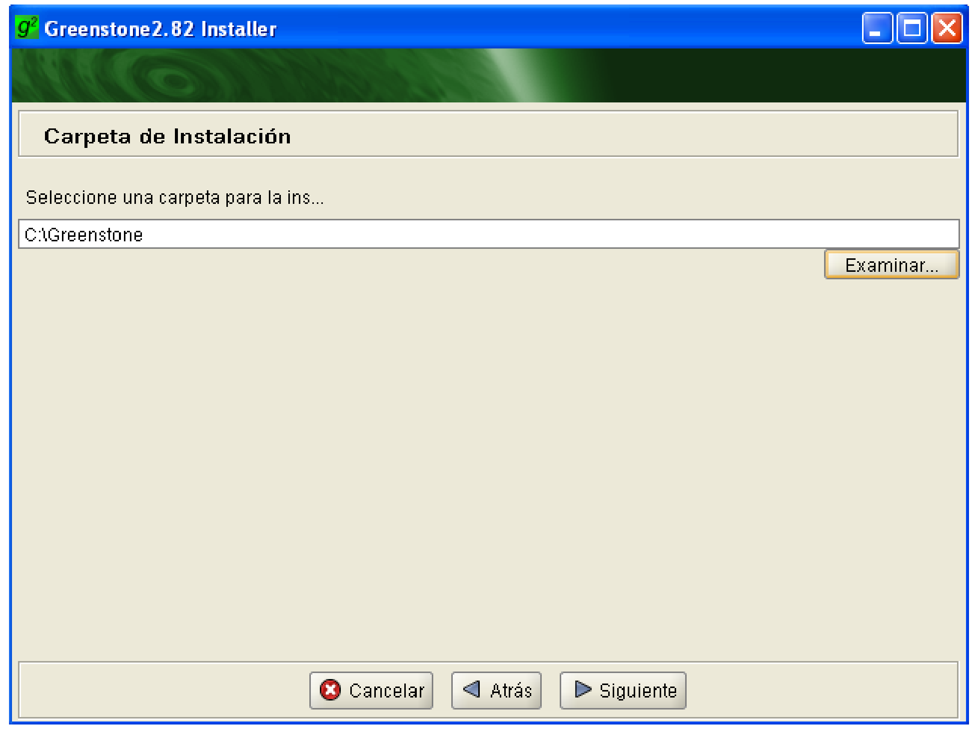


Figura 5.24. Selección del disco C:\ para instalar el programa Greenstone

En la próxima pantalla (Figura 5.25) se le consultará sobre los componentes del programa que desea instalar. Es recomendable seleccionarlos todos.



Figura 5.25. Pantalla para la elegir los componentes a instalar.

Recordemos que ImageMagicK es un paquete de software *open source* para la manipulación de imágenes. En la siguiente pantalla (Figura 5.26) se le cuestiona sobre su voluntad de incorporar mecanismos de control administrativos de modo que los usuarios no puedan modificar los contenidos de las colecciones. Se recomienda instalar los permisos administrativos y de esto hablaremos mas adelante.

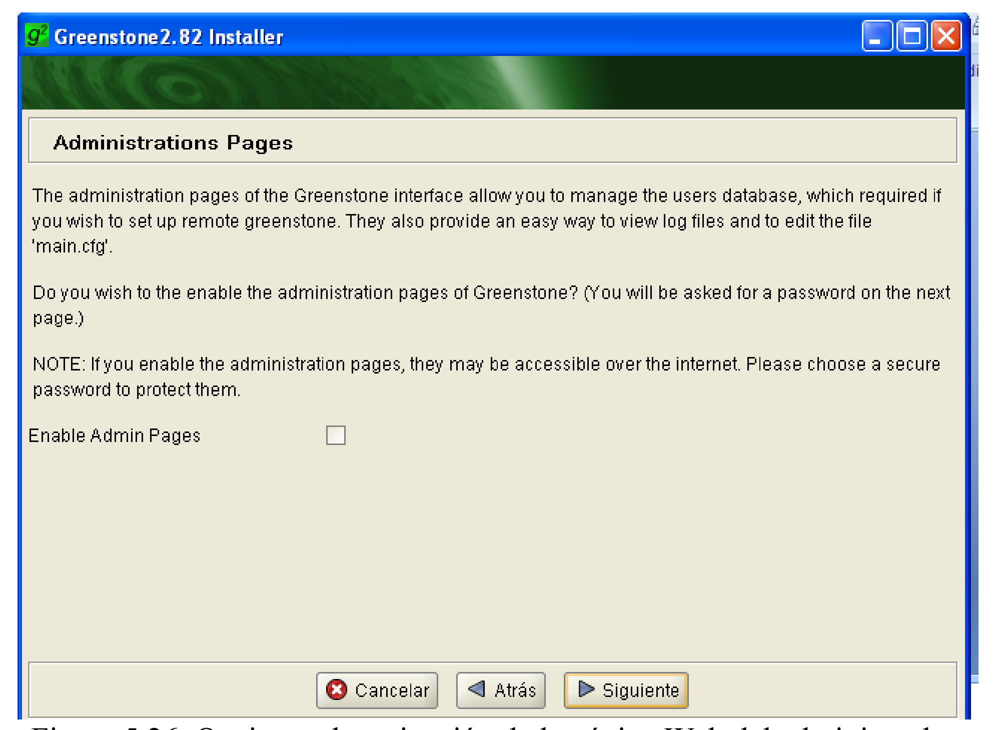


Figura 5.26. Opciones de activación de la página Web del administrador.

En la próxima pantalla (Figura 5.27) se le hace una consulta sobre su deseo o no de observar el progreso de la instalación.



Figura 5.27. Interfaz de consulta sobre si desea ver o no los detalles del proceso de instalación.

Si Ud. Accede a visualizar el proceso de instalación podrá ver una pantalla como la siguiente (Figura 5.28).

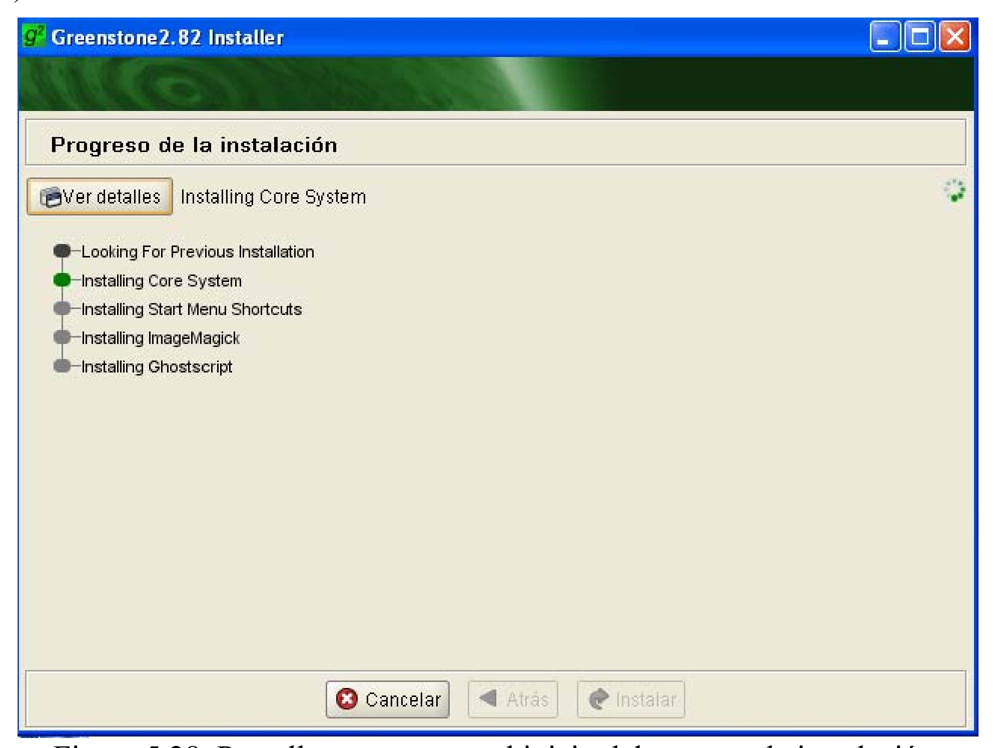


Figura 5.28. Pantalla que muestra el inicio del proceso de instalación.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

Esta pantalla permanecerá activa durante varios minutos y usted podrá observar el progreso de la instalación en las diferentes etapas. Una vez concluida la instalación aparecerá una ventana que le avisara que la instalación ha finalizado.

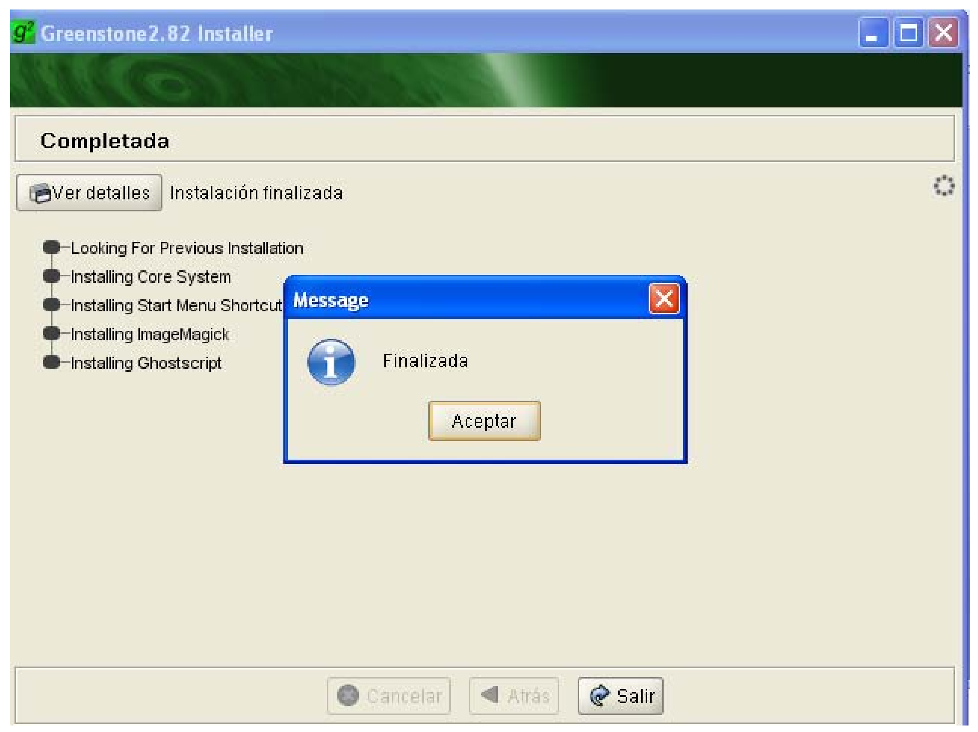


Figura 5.29. Pantalla con el mensaje que indica la terminación del proceso de instalación.

Y después de instalada la aplicación donde lo ha decidido (puede ser en disco D o C), esta será la apariencia de las carpetas que tiene el programa, si se dirige al interior del mismo.

Una vez concluida la instalación, para arrancar su sistema Greenstone pulse el botón Inicio, abra el menú Programas y seleccione Biblioteca Digital Greenstone (Greenstone Server). Esto hace aparecer un cuadro de diálogo: haga clic en Entrar en la biblioteca (Enter Library). (7)



Figura 5.30. Ejecución del Greenstone Server desde una instalación Windows.

Esto activa automáticamente su navegador Internet y carga la página principal de la Biblioteca Digital Greenstone. A partir de esta página se podrá tener acceso a la colección de demostración de Greenstone. El programa de la Biblioteca Local, le permite, a cualquier computadora equipada con Windows, distribuir colecciones Greenstone previamente creadas (7).

Todo el programa Greenstone se instala por defecto en el directorio /usr/local/gsdl si es el usuario root quien efectúa la instalación, y de otro modo en el directorio ~/gsdl (en que "~" es el directorio principal del usuario).

Ud. tiene además la opción para correr la GLI bajo Windows de explorar la carpeta gli que se encuentra en su instalación de Greenstone (por ejemplo usando el Explorador de Windows) y haciendo doble click en el archivo gli.bat. Este archivo verificará que Greenstone, el Ambiente de Ejecución Java y Perl están instalados e iniciará la Interfaz de Bibliotecario (7)

No queremos concluir el tema sin mencionar las características del archivo de configuración gsdlsite.cfg. El procedimiento básico de instalación crea un archivo de configuración de sitio genérico de Greenstone basado en las opciones de instalación que usted ha escogido. Para nuestra instalación este archivo es C:\Program Files\gsdl\cgi-bin\gsdlsite.cfg y su contenido es el siguiente (8):

```
# Site configuration file por Greenstone.
# Lines begining with
# are comments.
# This file should be placed in the same directory as your library
# executable file. It should be edited to suit your site.
# points to the GSDLHOME directory
gsdlhome "C:/Program Files/gsdl"
# this is the http address of GSDLHOME
# if your webservers DocumentRoot is set to $GSDLHOME
# then httpprefix can be commented out
httpprefix /gsdl
# this is the http address of the directory which
# contains the image for the interface.
httpimg /gsdl/images
# should contain the http address of this cgi script. This
# is not needed if the http server sets the environment variable
# SCRIPT_NAME
#gwcgi /cgi-bin/library
# maxrequests is the most requests a fastcgi process
# will serve before it exits. This can be set to a
# low figure (like 1) while debugging and then set
# to a high figure (like 10000) when everything is
# working well.
#maxrequests 10000
```

La Interfaz de Bibliotecario de Greenstone cuenta con una extensa variedad de opciones de ayuda en línea. Usted puede obtenerla haciendo click en el botón Ayuda que se encuentra en la parte superior derecha de la pantalla.

Para obtener mas ayuda sobre la instalación de Greenstone Ud. puede consultar también:

- 1-Biblioteca Digital Greenstone. Guía del Programador .Configuración del sitio Greenstone.

[Sitio Web en Internet] 2004. Disponible en:
<http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm>

- 2- Biblioteca Digital Greestone. Guía de Instalación. Capitulo 2.El procedimiento de instalación. [Sitio Web en Internet] 2007. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_installation_procedure.htm >-
- 3- Biblioteca Digital Greenstone. Guía de Instalación. Capitulo 5. Configuración de su sitio. [Sitio Web en Internet]. 2007. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm>

5.4- Consideraciones Finales

Greenstone es un programa realmente de gran complejidad por todas las utilidades y prestaciones que pueden obtenerse de el, pero esta programado de modo tal, que Ud. puede beneficiarse de las bondades del mismo sin tener que llegar a convertirse en un especialista en programación. Ud debe centrar su esfuerzo en comprender las instrucciones básicas para el uso del programa, disponer de la bibliografía necesaria y saber hacer uso de el, copiando las instrucciones para el uso del programa que tan generosamente ofrecen los diseñadores del mismo.

5.5- Referencias Bibliográficas

(1) Greenstone .org. Biblioteca Digital Greenstone. Guía del Programador .Configuración del sitio Greenstone. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm>

(2) Software de digitalizacion de publicaciones cientificas – Greenstone. Grupo CEDIPUR 35. Instituto de Planeamiento Urbano y Regional.Facultad de Arquitectura y Urbanismo – UNNE. [Libro en Internet]. 2006. Disponible en: <http://arq.unne.edu.ar/centros/institutos/ipur/publipur/greenstone.pdf >

(3) Pergamino Virtual.com. Definición de Servidor [Sitio Web en Internet]. 11/08/10. Disponible en: <http://www.pergaminovirtual.com/definicion/default.css>

(4) Greenstone.org. Guía de Instalación Greenstone. [Sitio Web en Internet]. 2002. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone>

(5) Greenstone.org. Biblioteca Digital Greenstone. Guía de Instalación. Capitulo 4. Instalación del Servidor Web. [Sitio Web en Internet]. 2007. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_setting_up_webserver.htm 2007>

(6) Wikipedia.org. Plataforma Java. [Sitio Web en Internet].28 de julio de 2011. Disponible en: <http://es.wikipedia.org/wiki/Plataforma_Java>

(7) Greenstone.org. Biblioteca Digital Greenstone. Guía de Instalación. Capitulo 2.El procedimiento de instalación [Sitio Web en Internet]. 2007. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_installation_procedure.htm>

(8) Greenstone.org. Biblioteca Digital Greenstone. Guía de Instalación. Capitulo 5. Configuración de su sitio. [Sitio Web en Internet]. 2007. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm>

(9) Greenstone.org. Ejercicio tutorial de Greenstone Actualización de una instalación de Greenstone

[Sitio Web en Internet]. 2007. Disponible en: <http://greenstone.sourceforge.net/wiki/index.php/Tutorial_exercises2010>

(10) Greenstone Digital Library Project LAB 1: Greenstone: Installing, browsing, building, University of Waikato New Zealand & Greenstone Support Network for South Asia. [Website in Internet]. 2010. Available in: <http://greenstonesupport.iimk.ac.in/greenstone2010-nepal/pdf/tutorial%20Exercish/LAB%201-Installing-%20browsing-%20building.pdf>

(11) Greenstone .org. Biblioteca Digital Greenstone. Guía del Programador .Configuración del sitio Greenstone. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm>

(12) Greenstone.org. Guía de Instalación Greenstone. [Sitio Web en Internet]. 2002. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone>

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

CAPÍTULO 6

6- Administración y gestión de usuarios en Greenstone

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH).

Palabras clave: Bibliotecas Digitales, Desarrollo Sostenible, Economía Digital, Estrategias de Inclusión Digital. , Sociedad de la Información, Software Libre, UNESCO.

6.1- Introducción

Para que Greenstone funcione correctamente, se deben instalar adecuadamente los permisos de acceso para determinados archivos (1). El procedimiento de instalación crea una configuración genérica basada en las opciones de instalación escogidas por usted; no obstante, su contenido puede personalizarse para hacer frente a distintas situaciones.

La función "de administración" establece los privilegios de cada tipo de usuario y de esto estaremos hablando en este capitulo.

6.2- El proceso de instalación del software

La instalación del software debe hacerse desde una cuenta con derechos para modificar determinados ficheros de configuración y permisos de ejecución (2).Destaque este aspecto en esta fase de la instalación como se observa en la Figura 6.1.

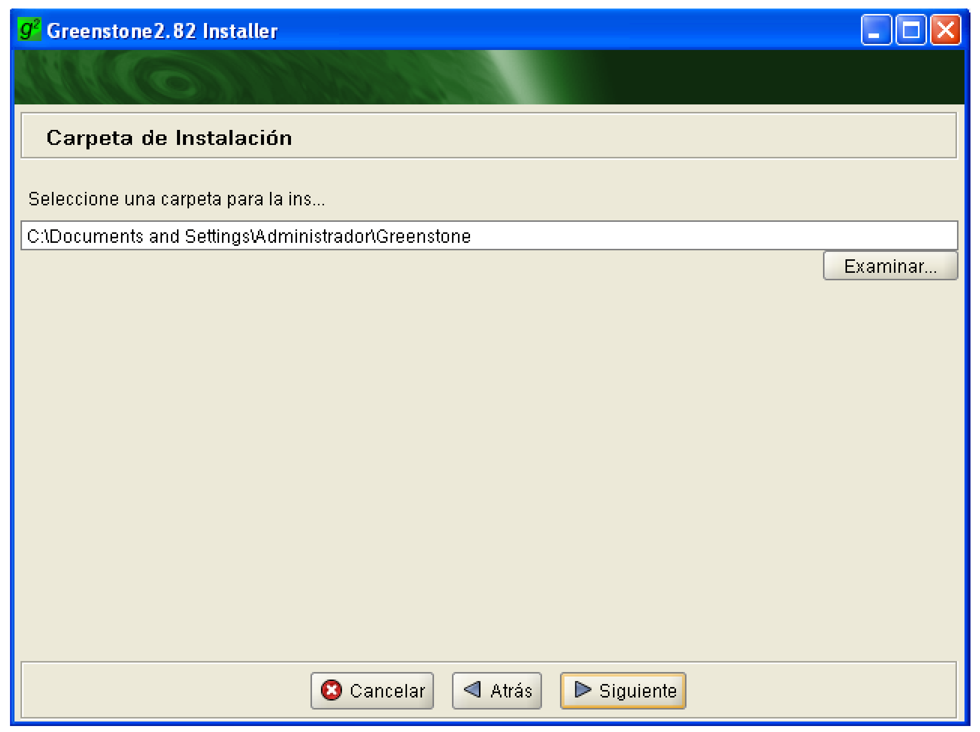


Figura 6.1-Instalación del software Greenstone en C:/ desde una cuenta de Administrador

En el curso del proceso de instalación que ya hemos revisado en el Capitulo 5 de este libro, existe un momento significativo en el cual Ud. debe precisar que tomará el control administrativo del software (Vea Figura 6.2).

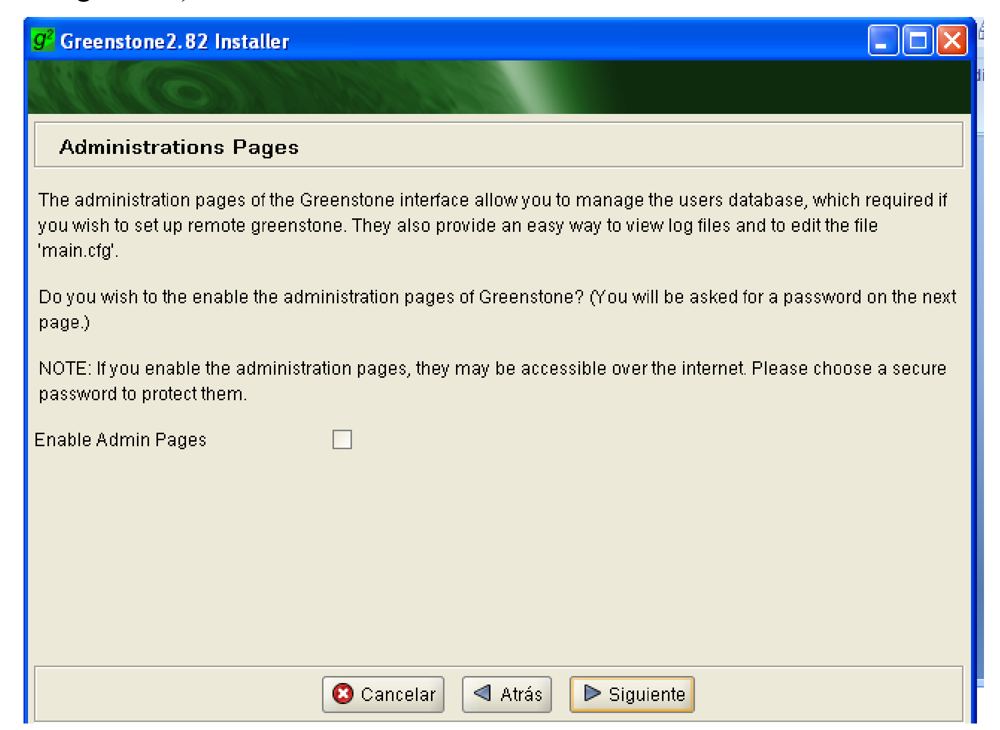


Figura 6.2- Determinación del control administrativo del software durante el proceso de instalación

En este momento Ud. debe indicar en la casilla correspondiente que asumirá el control administrativo.

El usuario "*administrador*" se llama admin, con una clave de acceso que usted escogió durante el proceso de instalación. El administrador está autorizado a añadir nuevos usuarios y a crear colecciones (3)

Debe usted saber que si la versión de biblioteca Web de Greenstone no se instala como se indica más arriba, cualquier persona podrá descargar cualquier archivo en la estructura de directorios gsdl. Esto comprende los archivos de índice y los documentos fuente de cualquiera de las colecciones que haya usted creado, la base de datos del usuario, los registros de utilización, etc.

Para evitar esto, puede usted reforzar fácilmente las medidas de seguridad modificando la configuración de su servidor Web. En el caso del servidor Web Apache, sustituya las líneas en el archivo de configuración indicadas en el apartado anterior por las siguientes (4):

```
Alias /gsdl/ “C:/Program Files/gsdl/”
<Directory C:/Program Files/gsdl>
Order allow, deny
Deny from all
 <FilesMatch “\. (gif|jpe?g|png|css|mov|mpeg|ps|pdf|doc|rtf|jar|class)$”>
Order allow, deny
Allow from all
</FilesMatch>
</Directory>
```

Esto significa que sólo se podrán descargar los archivos cuyas extensiones se ajustan a la expresión regular indicada en la línea FilesMatch.

Debemos recordar de los capítulos previos que existen dos archivos de configuración que controlan el funcionamiento de Greenstone: el archivo de configuración de sitio gsdlsite.cfg y el archivo de configuración principal main.cfg (5).

El archivo *gsdlsite.cfg* se utiliza para configurar el programa Greenstone para el sitio en que está instalado. Está concebido para registrar las opciones de configuración propias de un sitio determinado como, por ejemplo, el nombre del directorio en que se guarda el programa Greenstone, la dirección HTTP del sistema Greenstone, y si se está utilizando o no la función *fastcgi.*

El archivo *main.cfg* contiene información común a la interfaz de todas las colecciones presentadas en el sitio Greenstone. Incluye la dirección de correo electrónico del encargado del mantenimiento del sistema, indica si las páginas de estado y del Colector están activadas o no, si se llevan o no los registros de las actividades de los usuarios y si se utilizan o no " cookies " para identificar a los usuarios.

6.3- Permisos de acceso

Recordemos que el proceso de instalación comienza al preguntarnos donde instalar el software (3), pero Ud. debe verificar previamente si este software ha sido instalado con anterioridad en su computadora y si es así, dirigirse a la aplicación que posee el propio Greenstone para desinstalar el programa (Ver Figura 6.3), de modo que no quede nada del programa anterior y preferiblemente borrar todo vestigio de la aplicación anterior (e indicarlo así cuando en el proceso de desinstalación Ud. es interrogado sobre este aspecto) pues esto podría ocasionar dificultades en el futuro.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0



Figura 6.3-Icono que representa la opción de desinstalación del sistema Greenstone.

Para que Greenstone funcione correctamente, se deben instalar adecuadamente los permisos de acceso para determinados archivos (1).

Todos los archivos y directorios dentro de *C:\Program Files\gsdl* deben ser legibles por todos (o al menos legibles por el usuario de guiones CGI, que puede ser *"nobody"*). Para comprobar si los permisos de archivo están configurados correctamente, ejecute el programa *library.exe* a partir de la línea de comando. Si los archivos están en los lugares correctos pero los permisos están mal configurados, el programa arrancará desde la línea de comando –esto es, cuando usted la ejecuta- pero no desde un navegador –esto es, cuando el usuario *"nobody"* la ejecuta. Otra prueba es iniciar la sesión como otro usuario para ver si los permisos de archivo corresponden específicamente a su cuenta de usuario original (1).

Para funcionar desde un navegador Web, todos los directorios de Greenstone deben ser legibles por todos. Asimismo, el directorio *C:\Program Files\gsdl\etc* y todo su contenido deben ser modificables por cualquier persona. Éste es el directorio en el que el programa de biblioteca escribe el registro de uso, los registros de error y de inicialización y varias bases de datos del usuario. Si no desea usted que cualquier persona pueda modificar este directorio, puede establecer permisos para que el usuario de CGI sólo pueda modificar los archivos errout.txt, initout.txt, key.db, users.db, history.db y *usage.txt* (1).

Si los permisos de archivo no están configurados correctamente para C:\Program Files\gsdl\etc, es posible que no funcione la autenticación de usuario y el historial de búsqueda, y que no se genere un registro de uso (usage.txt) (1).

6.3.1- Autenticación

Para la autenticación se utiliza el sistema de cuentas de usuario de Greenstone. La información de los usuarios es almacenada en el archivo etc/users.db, y las páginas de Administración (llamadas desde la home de Greenstone) se usan para agregar, editar o borrar usuarios (6). Ver Figura 6.4.

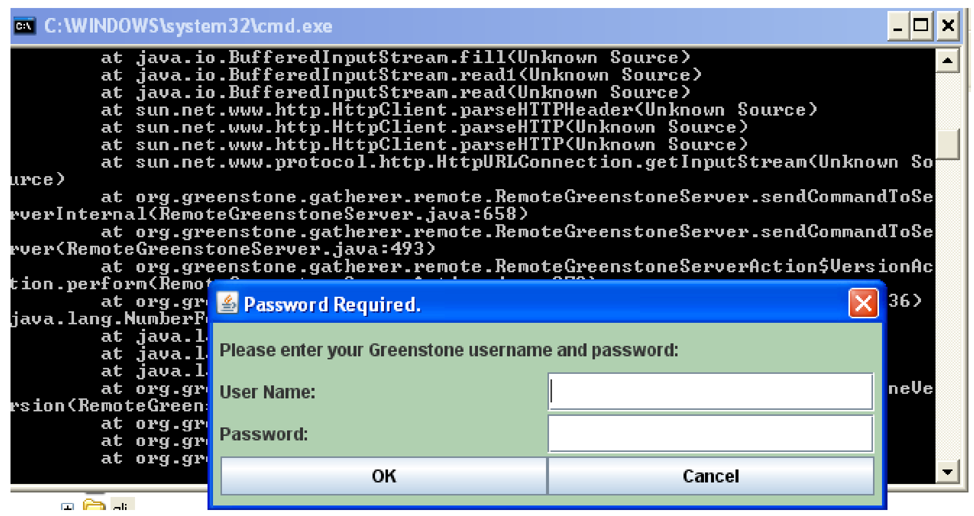


Figura 6.4.Ejemplo de la opción autentificación en el software Greenstone

También es posible determinar dentro del conjunto de colecciones de que Ud. dispone en su biblioteca cual será de acceso publico o no, mediante la opción "*La colección será de acceso publico*",tal y como muestra la Figura 6.5.



Figura 6.5.Opción determinar el acceso publico o no a la colección creada en la pestaña Formato

Ud. también puede seleccionar en que modo trabajar con la interfaz del bibliotecario (7) lo que conlleva por supuesto diferentes atributos y posibilidades dentro del sistema en la opción de preferencias de la ventana (Figura 6.6).

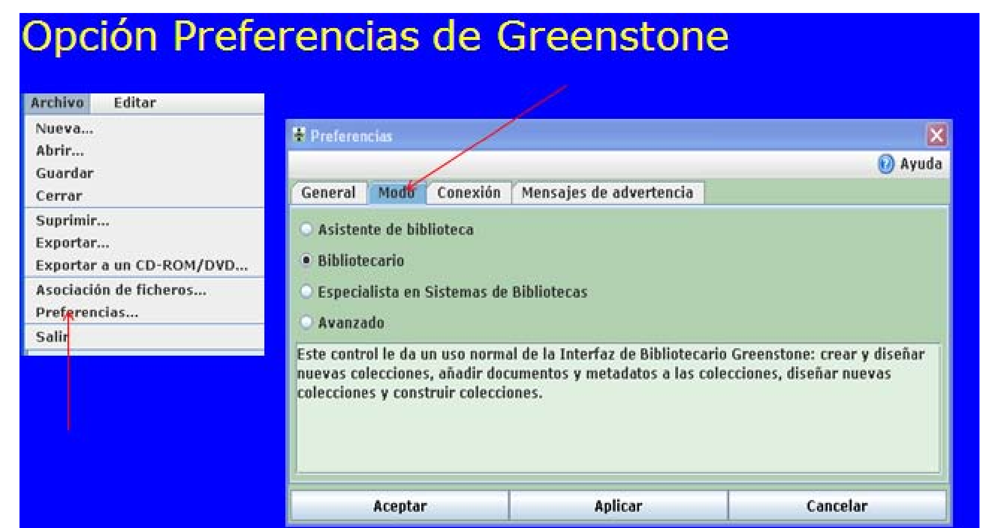


Figura 6.6. Opción Preferencias de Greenstone

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

Les invitamos a profundizar sobre el tema en: Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3.El Sistema de Ejecución de Greenstone. [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

6.4- Bondades del Sistema

6.4.1- Empleando Windows

Si quiere correr la interfaz del bibliotecario de Greenstone y esta utilizando Windows como sistema operativo Ud, puede dirigirse a a la carpeta gli del programa utilizando el Explorador de Windows y ejecutar un doble clic en el file gli.bat. Esto le permitirá corroborar que el ambiente Java y el Perl están correctamente instalados y comenzar el trabajo en esta inteface del bibliotecario (3).

6.4.2- La opción Ayuda

La interface del bibliotecario presenta un menú de ayuda que se hace visible mediante un botón en el top superior derecho. Ver Figura 6.7.

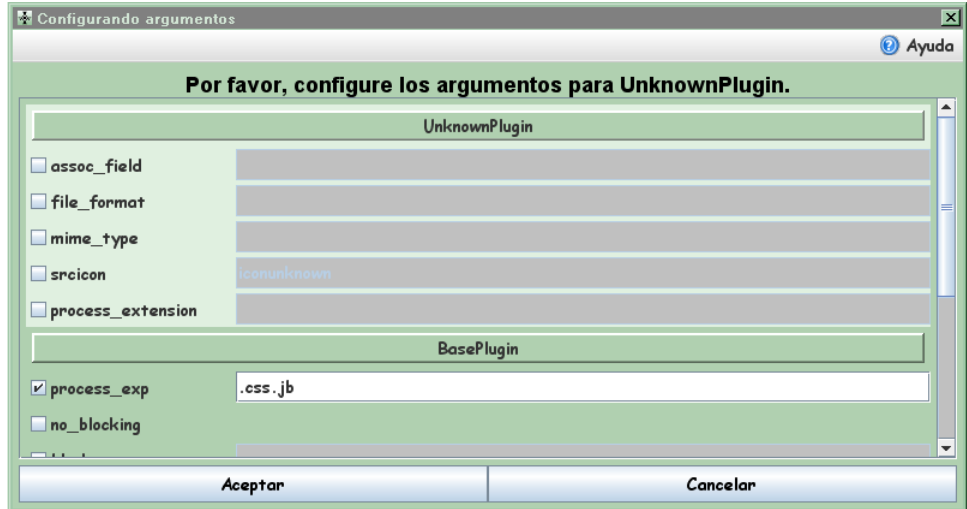


Figura 6.7. Identificación del menú de ayuda en todas las ventanas de trabajo del sistema Greenstone

6.4.3- Notificaciones por E-mail

El servidor puede configurarse para que le envíe un mail al administrador del sistema cada vez que finaliza el proceso de construcción de una colección. Para habilitar esta opción hay que editar el archivo "cgi-bin/gliserver.pl" y setear "$mail_enabled" en "1", y "$mail_to_address", "$mail_from_address", y "$mail_smtp_server" apropiadamente (6).

6.4.4- La opción "fast-cgi"

Esta opción acelera las ejecuciones de los guiones CGI manteniendo en memoria el ejecutable principal entre dos utilizaciones del programa, en vez de cargarlo a partir del disco cada vez que el programa Greenstone solicita una página Web. Se debe conseguir un equilibrio con la cantidad de memoria utilizada, que puede aumentar mientras más tiempo permanezca el programa en memoria (1).

Una vez generadas las páginas maxrequests, el programa CGI se cierra, liberando así el espacio de memoria ocupado. Para responder a la siguiente solicitud de página Web, el programa CGI se lee de nuevo a partir del disco duro y se inicia un nuevo ciclo de solicitudes de página. La mayoría de las

instalaciones utilizan el protocolo CGI estándar, lo que significa que el parámetro maxrequests puede ser ignorado sin problema (1).

Programas autónomos incluidos en Greenstone (7)	
setpasswd/	Ayuda para las contraseñas bajo Windows.
getpw/	Ayuda para las contraseñas bajo Unix.
txt2db/	Convierte un formato de texto ASCII de tipo XML en formato de base de datos de GNU.
db2txt/	Convierte el formato de base de datos de GNU en formato de texto ASCII de tipo XML.
phind/	Herramienta de consulta jerárquica en los grupos de palabras.
hashfile/	Calcula un identificador de documento único, basándose en el contenido del archivo.
mgpp/	Versión reescrita y actualizada del paquete *Managing Gigabytes* en C++
w32server/	Servidor de Biblioteca Local para Windows.
checkis/	Ayuda específica para la instalación de Greenstone bajo Windows.

6.5- Desventajas en Greenstone

6.5.1- Bloqueo de la colección

Cada colección puede ser abierta por una persona a la vez, para prevenir problemas de sincronización. Cuando se envía una petición al servidor para ejecutar una acción sobre la colección, el servidor chequeará la existencia del archivo gli.lck en la carpeta de la colección. Este archivo contiene el nombre de usuario de la persona que tiene bloqueada la colección. Cuando la colección se cierra, este archivo de bloqueo se borra (6).

Si la colección esta siendo bloqueada por una persona distinta a quien está haciendo la petición, entonces la acción falla. Esto es reportado al usuario y éste tiene la opción de “apropiarse” del bloqueo. Generalmente esto no es recomendado, ya que pueden perderse datos si hay múltiples usuarios editando la colección a la vez (6).

6.6- Mantenimiento del sitio y registro de utilización

Se pueden examinar tres tipos de registros: registros de uso, registros de errores y registros de inicialización. Los dos últimos sólo interesan de hecho a las personas encargadas del mantenimiento del programa (5).

Todas las actividades del usuario, es decir, cada página que el usuario visita, pueden quedar registradas en el programa Greenstone, aunque en los registros no se almacena información nominativa. La teneduría de registros (*logging*), desactivada por defecto, se activa incluyendo las líneas siguientes (5):

logcgiargs true
usecookies true

en el archivo principal de configuración del sistema. Ambas opciones tienen por defecto el valor false, por lo que es necesario asignarles el valor true para activar la teneduría de registros. La línea logcgiargs es la que activa o desactiva esta función. Activando *usecookies* se asigna a cada usuario un código de identificación único que permite seguir en el archivo de registro las interacciones de cada usuario (5).

Cada línea del registro de usuarios corresponde a una página visitada, y comprende: a) la dirección IP de la computadora del usuario, b) la marca de fecha y hora entre corchetes, c) los argumentos CGI entre paréntesis y d) el nombre del navegador del usuario (5).

Veamos a continuación resumidas en la tabla las opciones de configuración posibles (5).

Tabla 6.1. Archivos de macros del programa de interfaz de usuario

Opciones de configuración para el mantenimiento del sitio y el registro de utilización (*logging*)

	Valor	Finalidad
maintainer	NULL o una dirección de correo electrónico	Dirección de correo electrónico del encargado del mantenimiento del sitio que se utiliza para algunas notificaciones. Si es NULL se desactivan los eventos de correo electrónico.
MailServer	NULL o un nombre de servidor	Servidor de correo electrónico de salida para el sitio. Si es NULL, se utiliza *mail.dominio_del_encargado_del_mantenimiento* (por ejemplo, si el encargado del mantenimiento es *ayuda@ejemplo.com* el valor por defecto será *mail.ejemplo.com*.) Si ello no lleva a un servidor SMTP válido, los eventos de correo electrónico no funcionarán.
status	*enabled* (activado) o *disabled* (desactivado)	Determina si la página de “Mantenimiento y administración” debe estar disponible.
collector	*enabled* (activado) o *disabled* (desactivado)	Determina si la aplicación “Colector” de creación de colección para el usuario final debe estar disponible.
usecookies	*true* (verdadero) o *false* (falso)	Si el valor es *true* (verdadero), se acopia información relativa a los usuarios del sitio (mediante registros de identificación o *cookies*) que se guarda en *usage.txt* (sólo funciona si la opción *logcgiargs* tiene el valor *true*).
LogDateFormat	*LocalTime* (hora local) o *UTCTime* (hora universal) o *Absolute* (hora absoluta)	Formato en el que se consignan las fechas y las horas en los registros. *LocalTime* genera el formato “Thu Dec 07 12:34 NZDT 2000”, *UTCTime* es el mismo formato pero en tiempo universal (o GMT), y *absolute* es un número entero que representa el número de segundos transcurridos desde “00:00:00 01/01/1970 GMT”
LogEvents	*AllEvents* (todos los eventos) o *CollectorEvents* (eventos del Colector) o *disabled* (desactivado)	Registra determinados eventos en *events.txt*. *AllEvents* consigna todos los eventos de Greenstone, *CollectorEvents* sólo registra los eventos relacionados con el Colector, y *disabled* no registra ningún evento.
EmailEvents	*enabled* (activado) o *disabled* (desactivado)	Envía un mensaje por correo electrónico al encargado del mantenimiento (en caso de haberlo - véase la opción *mantainer*) cada vez que se produce un evento.
EmailUserEvents	*enabled* (activado) o *disabled* (desactivado)	Envía un mensaje por correo electrónico al usuario sobre determinados eventos (por ejemplo, cuando el Colector finaliza la creación de una colección).
macrofiles	lista de nombres de	Determina qué macros se encuentran a disposición

archivos de Macros del programa de interfaz de usuario.

6.7- Retirar Greenstone de un sistema Windows

Estos son los pasos aconsejables en el proceso de retirar completamente la versión existente de Greenstone antes de instalar una nueva versión (8):

1. Asegúrese que no se está ejecutando Greenstone
2. Retire la versión anterior yendo al panel de control de Windows. Haga clic en Añadir o retirar Programas,
3. Escoja Software de Greenstone de Biblioteca digital, y Retirar . (Para hacerlo, usted puede necesitar privilegios del administrador de Windows.)
4. Al final de este procedimiento se le preguntará si quiere retirar todas sus colecciones, debe contestar No si quiere preservar su trabajo.

Pueden aparecer problemas si no se han retirado completamente las versiones anteriores de Greenstone. Para limpiar su sistema, mueva su carpeta local *collect* de Greenstone que contiene todas las colecciones, al *desktop*. Después busque la carpeta C:\Program *Files\gsdl* o *C:\ProgramFiles\Greenstone*, que es donde se instala usualmente Greenstone y si existe, retírela totalmente (8).

6.8- Consideraciones Finales

Hemos considerado todas las instrucciones precisas para lograr trabajar este software con la mayor eficiencia posible ya que de no contemplar los diferentes aspectos relacionados, podrían enfrentar dificultades en la implementación del programa.

6.9- Referencias Bibliográficas

(1) Greenstone.org. Biblioteca Digital Greenstone. Guía de Instalación. Capitulo 5. Configuración de su sitio. [Sitio Web en Internet]. 2007. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm>

(2) Tramullas J. Biblioteca Digital Greenstone. Capitulo 10. En.Software Libre para Servicios de Información [Libro en Internet] Digital. Madrid: Pearson Prentiss Hall, 2006. Disponible en:<http://tramullas.com>

(3) Greenstone.org. Biblioteca Digital Greenstone. Guía de Instalación. Capitulo 2.El procedimiento de instalación [Sitio Web en Internet]. 2007. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_installation_procedure.htm>

(4) Greenstone.org. Biblioteca Digital Greenstone. Guía de Instalación. Capitulo 4.Instalación del servidor Web. [Sitio Web en Internet]. 2007. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_installation_procedure.htm >

(5) Greenstone.org. Biblioteca Digital Greenstone. Guía del Usuario, Capitulo 4. Administración [Sitio Web en Internet].2007. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_administration.htm>

(6) Wiki.greenstone. Remote Greenstone. [Website in Internet].4 october 2010. Available from: <http://wiki.greenstone.org/wiki/index.php/Remote_Greenstone>

(7) Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3.El Sistema de Ejecución de Greenstone. [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

(8) Greenstone.org. Ejercicio tutorial de Greenstone Actualización de una instalación de Greenstone [Sitio Web en Internet]. 2007. Disponible en: <http://greenstone.sourceforge.net/wiki/index.php/Tutorial_exercises>

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

CAPÍTULO 7

7- La arquitectura interna de Greenstone: usabilidad del sistema

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH)

Palabras clave: Bibliotecas Digitales, Desarrollo Sostenible, Economía Digital, Estrategias de Inclusión Digital. , Sociedad de la Información, Software Libre, UNESCO

7.1- Introducción

El término "*Arquitectura de la Información*" (AI) fue utilizado por primera vez por Richard Saul Wurman en 1975, quién la define como (1):

> *"El estudio de la organización de la información con el objetivo de permitir al usuario encontrar su vía de navegación hacia el conocimiento y la comprensión de la información".*

Hemos querido subrayar este concepto aquí, ya que lo que pretendemos en este capitulo es mostrar las interioridades del programa que justifican una funcionalidad dada.

7.2- La distribución de carpetas en el software Greenstone. El lenguaje propio de Greenstone basado en macros

7.2.1- El característico Lenguaje de Macros de Greenstone

El lenguaje de macros, se utiliza para conferir al sistema de Biblioteca Digital Greenstone un estilo coherente y crear interfaces en diferentes lenguas. La interacción con la biblioteca genera la estructura básica de las páginas Web y las macros del directorio GSDLHOME/macros aportan el contenido.

En Greenstone no hay pagina HTML para mostrar la interfaz pues lo que tiene es un lenguaje propio basado en macros. Las Macros definen tanto las cadenas de formato de las interfaces como las variables utilizadas en Código C++ .Utilizando macros el usuario puede fácilmente personalizar la interface de usuario sin cambiar nada del código C++ (2) como veremos en la segunda parte de este libro.

7.2.1.1- ¿Donde se encuentra el directorio GSDLHOME?

Si nos dirigimos a la carpeta donde hemos instalado el software (directorio GSDLHOME) esta seria el conjunto de carpetas y elementos que encontraríamos. Recuerden que el objetivo de este capitulo es hacer una correspondencia entre los elementos físicos que caracterizan al software y las funcionalidades que nos ofrece.

Esta seria la apariencia del contenido del directorio Gereenstone que hemos instalado en la

computadora (Figura 7.1).

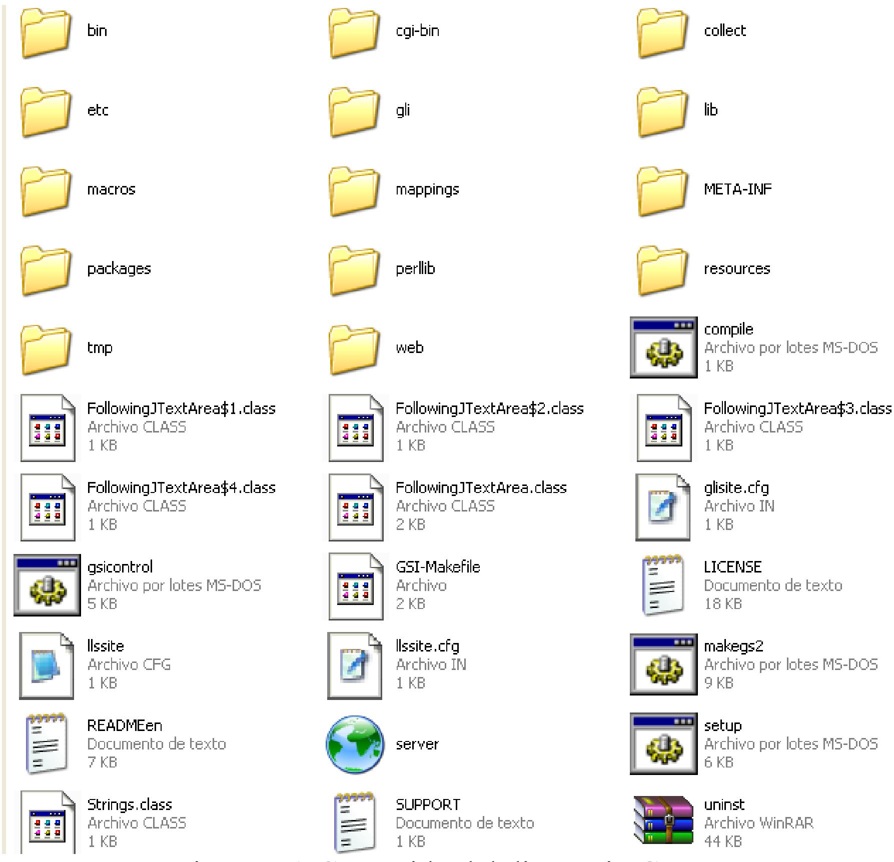


Figura 7.1-Contenido del directorio Greenstone.

La tercera carpeta en la primera columna (Figura 7.2) contiene las macros de visualización. Existen también las llamadas macros de configuración y de estas hablaremos posteriormente.

about amharic amharic2 arabic arabic2 armenian armenian2 authen base
bengali browse bsummary bulgarian catalan catalan2 chinese chinese2 chinese-trad

japanese japanese2 kannada kazakh kazakh2 kirghiz languages latvian latvian2
malayalam malayalam2 maori marathi marathi2 mongolian mongolian2 nav_css nav_ns4
nepali nzdlhome polish port-br port-pt pref pushto query romanian
romanian2 russian russian2 serbian sinhalese slovak slovak2 spanish spanish2
status style tamil telugu thai thai2 tip turkish turkish2
ukrainian urdu urdu2 usability users vietnamese vietnamese2 yourhome yourhome-ar
yourhome-es yourhome-fr

Figura 7.2. Contenido de la carpeta macros

Por tanto la carpeta macros es de gran importancia ya que Greenstone se basa en el trabajo de estos elementos para asumirlos como lenguaje propio (2). A una macro se le reconoce por su estructura, su nombre empieza y termina con un guion bajo y encierra su contenido entre llaves como se muestra en el siguiente ejemplo :_nombre de macro_{contenido}

Siguiendo una línea horizontal imaginaria estas serian las tres carpetas que aparecen dispuestas en el conjunto del contenido del directorio Greenstone (Figura 7.3)

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0.

bin	Código ejecutable
cgi-bin	Guiones CGI
collect	Colecciones Greenstone (comprende la colección de demostración, los índices creados previamente y los documentos fuente).

Figura 7.3. Primeras tres subcarpetas del directorio Greenstone

A continuación veamos los contenidos de cada una de ellas. Les proponemos que de modo práctico sigan por ustedes mismos la ejecución de estos pasos y esto los conducirá ciertamente a un mejor dominio del software.

En orden de aparición inmediata como deben haber destacado se encuentra la carpeta bin . En esta carpeta se encuentran alojados a su vez los siguientes elementos que se muestran a 'continuación (Figura 7.4).

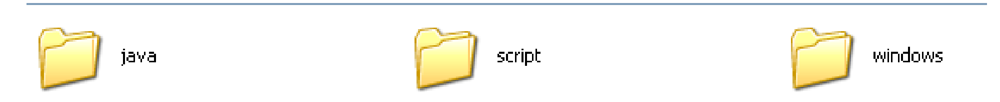


Figura 7.4. Contenido de la carpeta bin.

La carpeta bin contiene los códigos ejecutables por ello resultará de mucho interés introducirnos en ella para conocer las interioridades del programa. Así si hacemos doble clic sobre la carpeta java, podemos visualizar varias subcarpetas como se indica a continuación (Figura 7.5).



Figura 7.5. Contenido de la carpeta java.

Aquí vale la pena destacar lo siguiente:

- Lucene es uno de los buscadores que utiliza Greenstone , y hablaremos de el en el capitulo sobre Sistema de Búsqueda en Greenstone
- Xalan es el paquete que contiene el programa apache para el Servidor Web
- y recuerden que el GLI es la interfaz gráfica del software o ***G**reenstone **L**ibrary **I**nterfaz*

Si hacemos doble clic sobre el fichero compactado que indica GLIserver, observaremos los elementos que destacan allí (Figura 7.6):

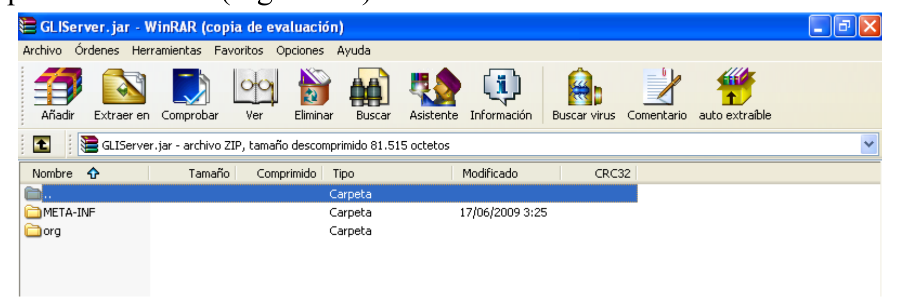


Figura 7.6. Contenido de la carpeta comprimida con WinRar, GLIserver

Si pinchamos en la carpeta org contenida en GLI server como se muestra en la figura anterior, viajaremos por las carpetas Gereenstone, gatherer y remote hasta encontrar finalmente las

instrucciones que garantizan el trabajo de esta interface (Figura 7.7).

greenstone gatherer remote

ActionQueue.class	3.728	1.838	Fichero class	17/06/2009 15:25	CF419B55
RemoteGreensto...	279	175	Fichero class	17/06/2009 15:25	743DAE22
RemoteGreensto...	1.387	774	Fichero class	17/06/2009 15:25	DF11A3C8
RemoteGreensto...	1.678	650	Fichero class	17/06/2009 15:25	791AC4E7
RemoteGreensto...	1.022	427	Fichero class	17/06/2009 15:25	8A5F1024
RemoteGreensto...	17.116	6.989	Fichero class	17/06/2009 15:25	C2459431
RemoteGreensto...	412	247	Fichero class	17/06/2009 15:25	53B11195
RemoteGreensto...	1.485	681	Fichero class	17/06/2009 15:25	7121267A
RemoteGreensto...	2.191	1.013	Fichero class	17/06/2009 15:25	DC5AF4BD
RemoteGreensto...	1.989	917	Fichero class	17/06/2009 15:25	301D4091
RemoteGreensto...	1.998	937	Fichero class	17/06/2009 15:25	94B7F721
RemoteGreensto...	1.768	816	Fichero class	17/06/2009 15:25	479C0CA9
RemoteGreensto...	2.302	1.078	Fichero class	17/06/2009 15:25	CDE1855D
RemoteGreensto...	1.388	711	Fichero class	17/06/2009 15:25	1BA2BE62
RemoteGreensto...	1.933	937	Fichero class	17/06/2009 15:25	ABD10D8B
RemoteGreensto...	1.630	766	Fichero class	17/06/2009 15:25	09B241B3
RemoteGreensto...	1.093	516	Fichero class	17/06/2009 15:25	CD9D43B1
RemoteGreensto...	824	426	Fichero class	17/06/2009 15:25	8BDD49F6
RemoteGreensto...	2.363	1.099	Fichero class	17/06/2009 15:25	5A456EFF
RemoteGreensto...	2.086	967	Fichero class	17/06/2009 15:25	A0C52DAD
RemoteGreensto...	1.901	861	Fichero class	17/06/2009 15:25	66ACBD1C
RemoteGreensto...	2.337	1.073	Fichero class	17/06/2009 15:25	688077D4
RemoteGreensto...	3.071	1.454	Fichero class	17/06/2009 15:25	83CDD366
RemoteGreensto...	809	408	Fichero class	17/06/2009 15:25	5AAA205F
RemoteGreensto...	3.296	928	Fichero class	17/06/2009 15:25	6FC89D42
Unzip.class	2.308	1.326	Fichero class	17/06/2009 15:25	1142964A
ZipCollectionArchi...	2.206	1.143	Fichero class	17/06/2009 15:25	BC991EEB
ZipCollectionArchi...	1.641	850	Fichero class	17/06/2009 15:25	AFA67B48
ZipCollectionConf...	294	184	Fichero class	17/06/2009 15:25	814C345D
ZipCollectionConf...	1.240	594	Fichero class	17/06/2009 15:25	D990B5D7
ZipCollectionConf...	802	371	Fichero class	17/06/2009 15:25	A0704738
ZipCollectionConf...	3.979	1.910	Fichero class	17/06/2009 15:25	E7C79211

Figura 7.7. Contenido de la carpeta org

Les dejamos como trabajo independiente la idea de reconocer los elementos que contiene la carpeta *script* y pasemos a continuación a la carpeta Windows adelantándoles que la carpeta gimp al interior de la carpeta *script* contiene los elementos que controlan la interfaz gráfica, como son los botones, flash etc

Si abrimos entonces el contenido de la carpeta Windows de bin (Figura 7.8) observaremos lo siguiente:

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0.

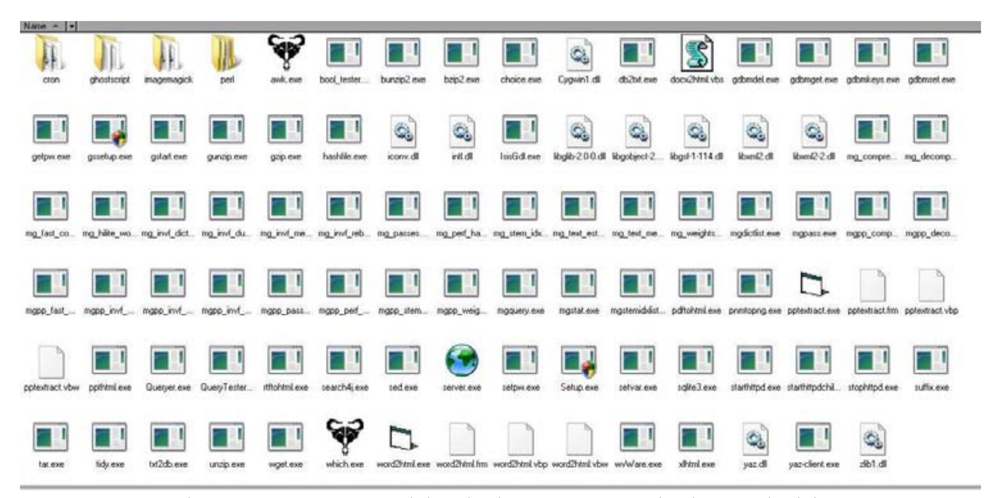


Figura 7.8. Contenido de la carpeta Windows de bin

¿Que es lo que vale la pena señalar en esta carpeta? Varias cosas:

1. El programa *Ghostscript* asociado al programa Greenstone del que ya hemos hablado en el capitulo anterior (Figura 7.9)

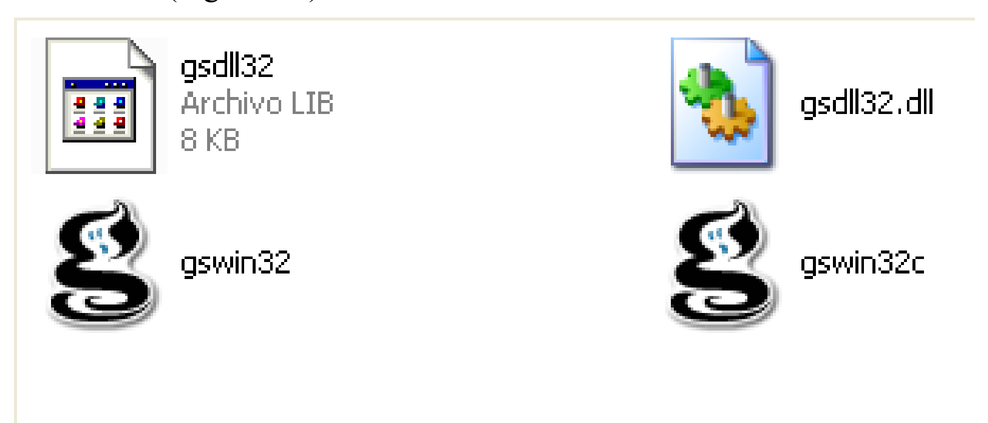


Figura 7.9.Contenido de la carpeta *Ghostscript*

2. El ImageMagick (Figura 7.10) también asociado al trabajo de Greenstone del cual también ya hemos hablado y que deben recordar de la clase de instalación del Greenstone cuando el proceso requería especialmente de su instalación como también sucedió con *Ghostscript*

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0.

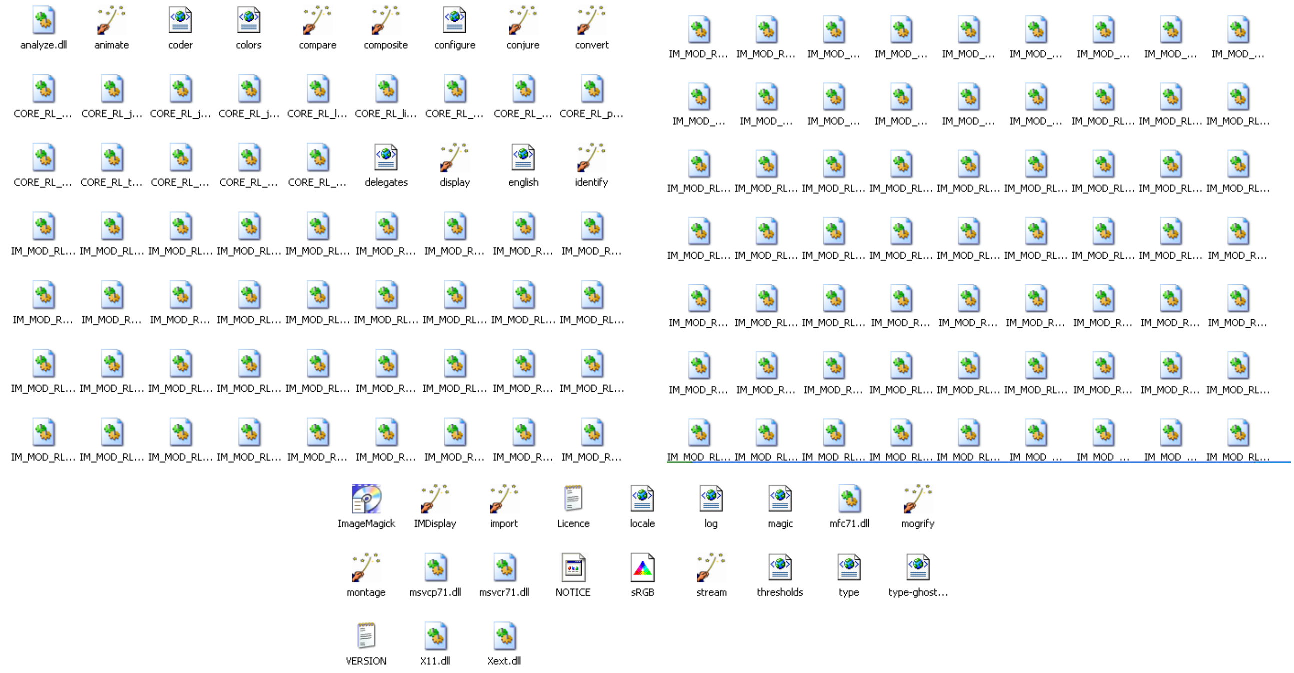


Figura 7.10. Contenido de la carpeta ImageMagick

3. La carpeta *perl* (Figura 7.11) que es un lenguaje asociado al Greenstone, como les habíamos mencionado en el capítulo sobre los atributos y condiciones de trabajo del software Greenstone

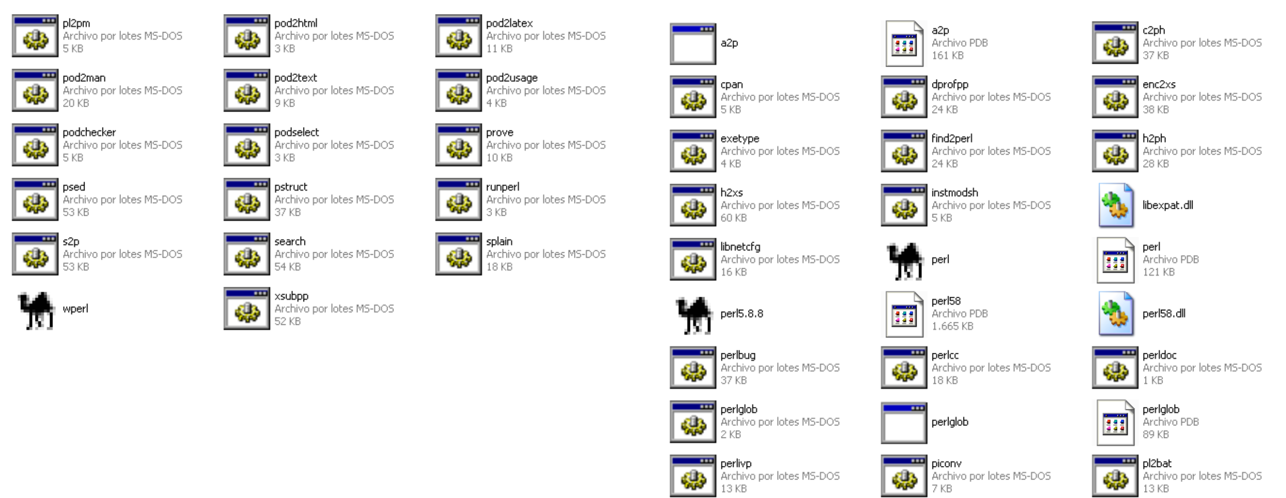


Figura 7.11. Contenido de la carpeta Perl.

Tenemos por ejemplo que los conectores o pluggins que ya hemos mencionado previamente se escriben en lenguaje Perl y que los clasificadores que nos permitirán una adecuada recuperación también son objetos Perl.

Si seguimos analizando el contenido de las carpetas según el orden imaginario que hemos trazado hallaríamos ahora la carpeta cgi-bin, este es el contenido de la carpeta que destaca por contener los guiones cgi (Figura 7.12). cgi-bin es un directorio desde el cual el servidor Web trata documentos como programas ejecutables. La directiva ScriptAlias de Apache se utiliza para crear un directorio cgi-bin. Cabe señalar que esta directiva puede hacer de cualquier directorio un directorio CGI ejecutable, ¡sin que se llame forzosamente “cgi-bin”! A la inversa, un directorio llamado “cgi-bin” no tiene nada especial si no se le ha aplicado la directiva ScriptAlias (3).

Recordemos que un Guión CGI es un Código asociado a un botón, un menú, o un enlace en una página Web, que especifica lo que ha de hacer la computadora central cuando se hace clic en ese objeto.

Algo importante es que los cgi-bin deben tener una extensión especial .cgi, o bien, encontrarse en un directorio especial llamado cgi-bin. Cuando se referencie en una pagina html un binario, cualquier fichero con la extensión cgi, en el servidor se ejecutará este programa y la salida se enviará al cliente de la WWW (4).

Si continuamos nuestro viaje imaginario por las carpetas de Greenstone veremos que al interior de la carpeta demo perteneciente a collect (Figura 7.12) se encuentran varios archivos, por su importancia revisemos sus características.

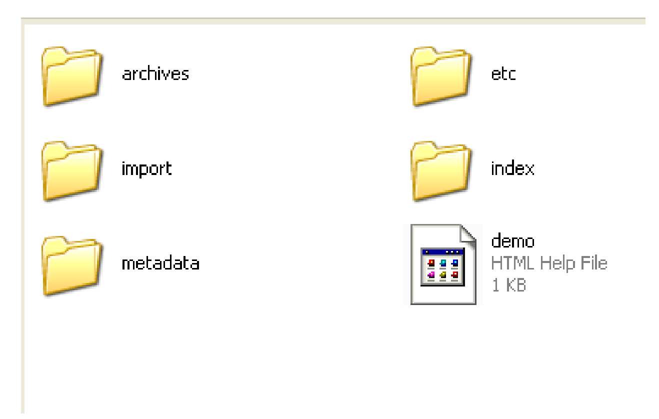


Figura 7.12. Contenido carpeta demo

La primera sub carpeta archivos contiene los Hash de los documentos integrados a las colecciones (Figura 7.13).

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0.

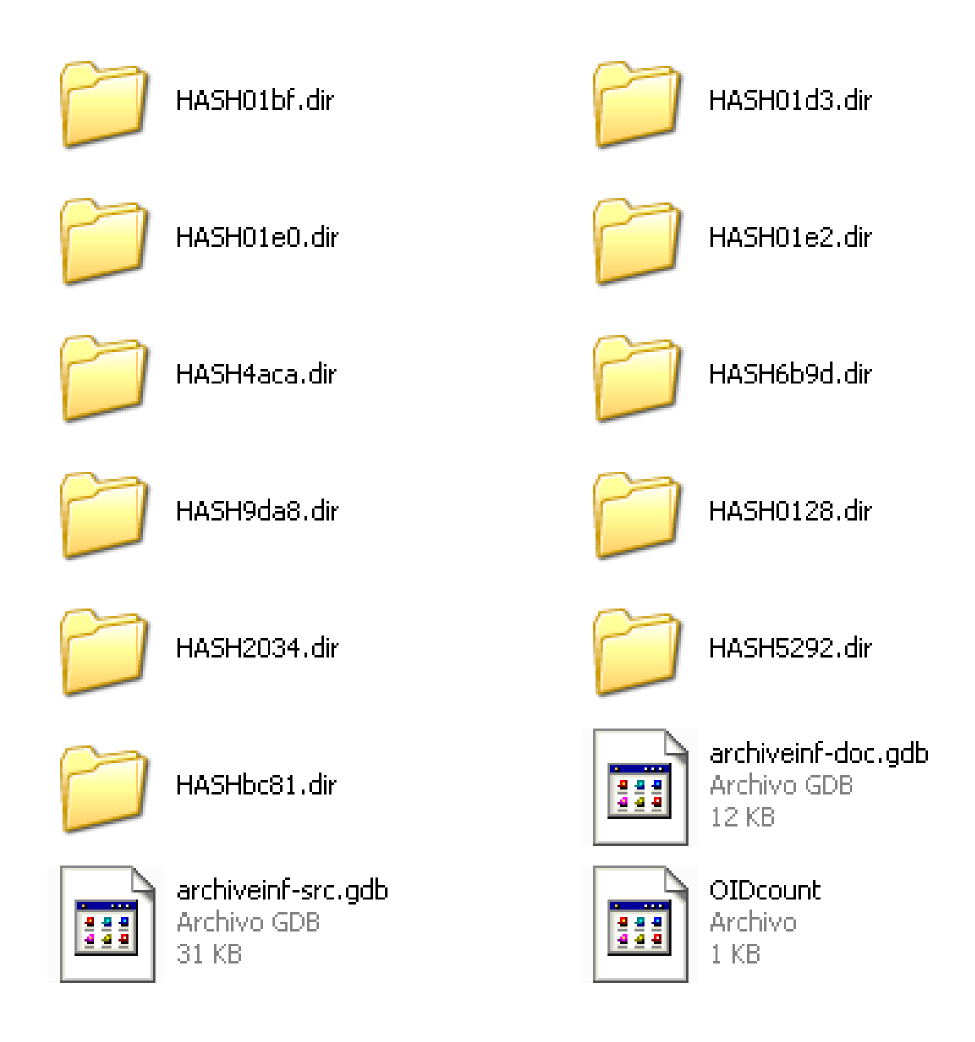


Figura 7.13. Contenido carpeta archivos

7.2.2- ¿Que son los Hash?

El Hash es como un Identificador único que cada sección del documento tiene asociado un hasla es un algoritmo único para cada documento o parte del documento (3).

Cada parte del documento tiene asociado un hash que es un algoritmo único como ya hemos señalado para cada documento o parte del documento. El sistema asigna un sistema alfanumérico único que nunca se podrá utilizar aunque se elimine de forma lógica

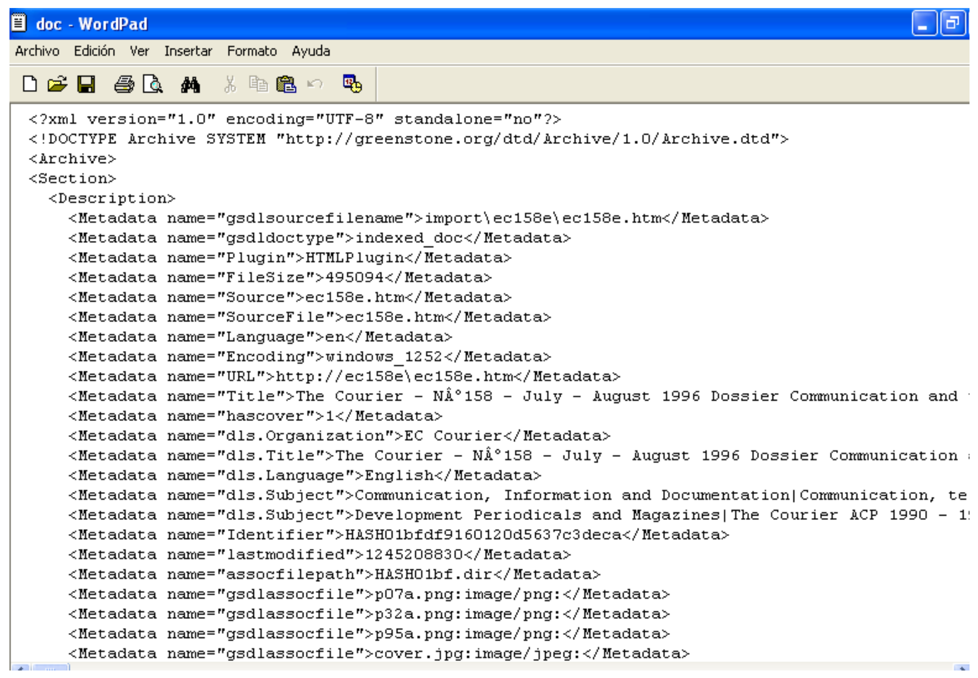
Se trata de un programa autónomo incluido en Greestone que calcula un identificador de documento único, basándose en el contenido del archivo. El sistema asigna un sistema alfanumérico único que nunca se podrá utilizar aunque se elimine de forma lógica. A continuación les mostramos los detalles internos de un Hash (Figura 7.14)

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0.



Figura 7.14. Contenido del Hash de un documento

Si abrimos con el programa word-pad el fichero doc. Esta es la imagen que obtendremos (Figura 7.15).




```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE Archive SYSTEM "http://greenstone.org/dtd/Archive/1.0/Archive.dtd">
<Archive>
<Section>
  <Description>
    <Metadata name="gsdlsourcefilename">import\ec158e\ec158e.htm</Metadata>
    <Metadata name="gsdldoctype">indexed_doc</Metadata>
    <Metadata name="Plugin">HTMLPlugin</Metadata>
    <Metadata name="FileSize">495094</Metadata>
    <Metadata name="Source">ec158e.htm</Metadata>
    <Metadata name="SourceFile">ec158e.htm</Metadata>
    <Metadata name="Language">en</Metadata>
    <Metadata name="Encoding">windows_1252</Metadata>
    <Metadata name="URL">http://ec158e\ec158e.htm</Metadata>
    <Metadata name="Title">The Courier - NÂ°158 - July - August 1996 Dossier Communication and
    <Metadata name="hascover">1</Metadata>
    <Metadata name="dls.Organization">EC Courier</Metadata>
    <Metadata name="dls.Title">The Courier - NÂ°158 - July - August 1996 Dossier Communication
    <Metadata name="dls.Language">English</Metadata>
    <Metadata name="dls.Subject">Communication, Information and Documentation|Communication, te
    <Metadata name="dls.Subject">Development Periodicals and Magazines|The Courier ACP 1990 - 1
    <Metadata name="Identifier">HASH01bfdf9160120d5637c3deca</Metadata>
    <Metadata name="lastmodified">1245208830</Metadata>
    <Metadata name="assocfilepath">HASH01bf.dir</Metadata>
    <Metadata name="gsdlassocfile">p07a.png:image/png:</Metadata>
    <Metadata name="gsdlassocfile">p32a.png:image/png:</Metadata>
    <Metadata name="gsdlassocfile">p95a.png:image/png:</Metadata>
    <Metadata name="gsdlassocfile">cover.jpg:image/jpeg:</Metadata>
```

Figura 7.15. Contenido del documento doc. del Hashh01bf.

Si abrimos el archivo doc. que vimos previamente de la carpeta Hash01bf , pero esta vez con Word de Office entonces visualizamos la cadena de procesos que se producen en la asignación de metadatos y nos resulta muy fácil visualizar la asignación del pluggin adecuado de html que requiere este documento en particular (Figura 7.16).

Archive
Section
Description
Metadata import\ec158e\ec158e.htm Metadata
Metadata indexed_doc Metadata
Metadata HTMLPlugin Metadata
Metadata 495094 Metadata
Metadata ec158e.htm Metadata
Metadata ec158e.htm Metadata
Metadata en Metadata
Metadata windows_1252 Metadata

Figura 7.16. Contenido del documento doc. del Hashh01bf abierto con Microsoft Word

En la carpeta collect (la ultima de las tres primeras carpetas que hemos estado analizando) se hallan las subcarpetas demo y Modelcoll, o Molde de plantilla en el cual se generan nuevas colecciones (Figura 7.17).

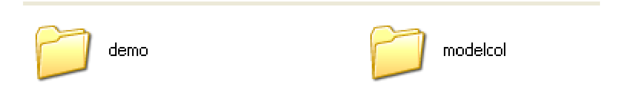


Figura 7.17. Contenido de la carpeta collect

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0.

Las siguientes tres carpetas en orden de aparición son; etc, gli e images

1. **etc** Archivos de configuración, registros, etc.
2. **gli** herramienta de la Interfaz de la Biblioteca Digital Greenstone
3. **images** Imágenes utilizadas por la interfaz

Figura 7.18. Siguientes tres subcarpetas del directorio Greenstone

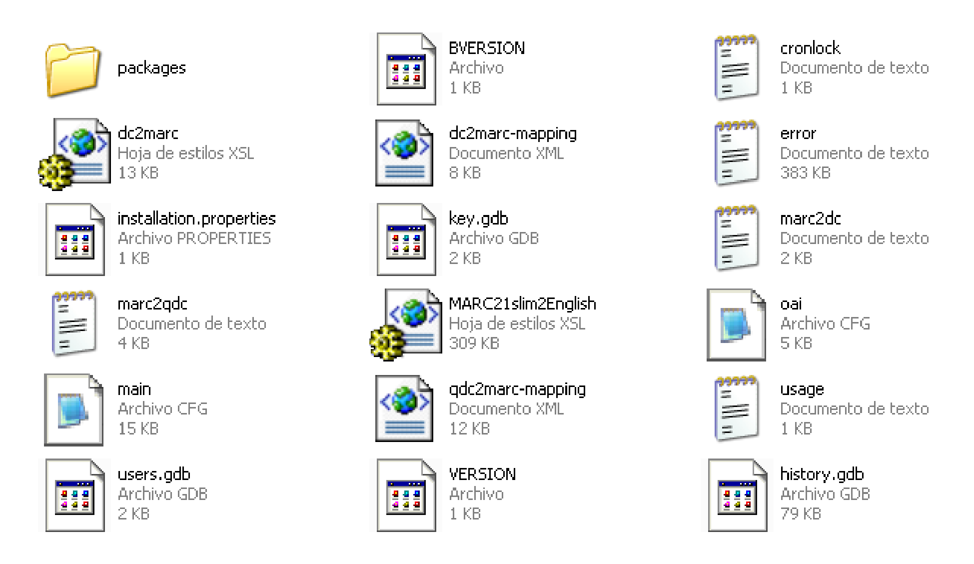


Figura 7.19. Contenido de la carpeta etc

El contenido de la carpeta etc se muestra en la Figura 7.19. Es importante señalar que en la carpeta etc encontraremos *main.cfg* (el quinto icono de la primera fila) y que este archivo a su vez tiene configuraciones esenciales para el sistema.

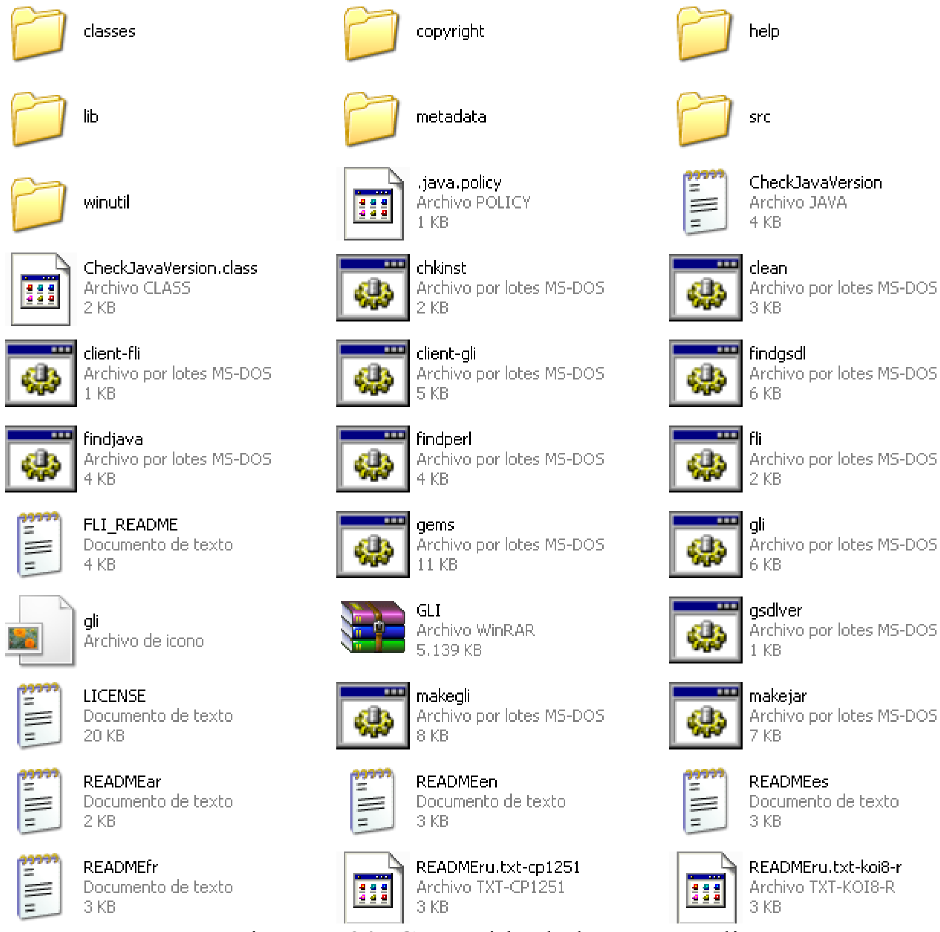


Figura 7.20. Contenido de la carpeta gli

El contenido de la carpeta gli se muestra en la Figura 7.20. La carpeta lib que sigue a gli aloja todo lo concerniente a la ejecución del programa java. Y en esta carpeta se encuentran alojados varios elementos que tributan a las características del servidor, al *gdsl* y a la propia lógica interna del sistema.

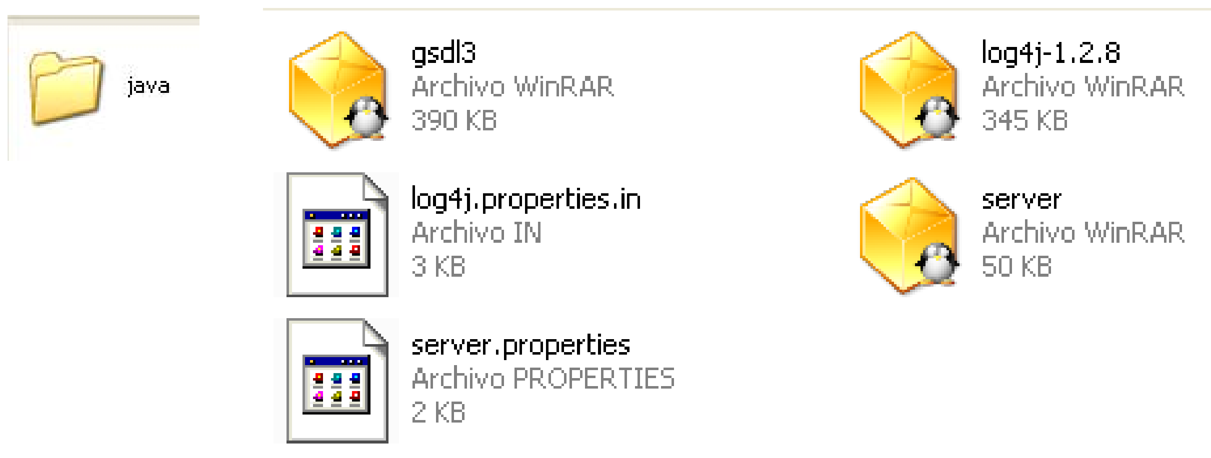


Figura 7.21. Contenido de la carpeta java de lib.

macros	Macros de visualización
perllib	Módulos Perl utilizados para la creación
mappings	Tablas de traducción Unicode

Figura 7.22. Siguientes tres subcarpetas del directorio Greenstone.

En nuestra lógica imaginaria este grupo de carpetas representarían la tercera línea en orden de aparición.

De esta carpeta de macros ya hemos hablado al inicio del capitulo. Debemos apuntar que existen macros con comportamiento predefinido y macros que definen su acción mediante el contenido que se les asigne. Los archivos de macros tienen la extensión .dm. En la carpeta *perlib* de Greenstone se encuentra la carpeta *pluggins* (Figura 7.23).

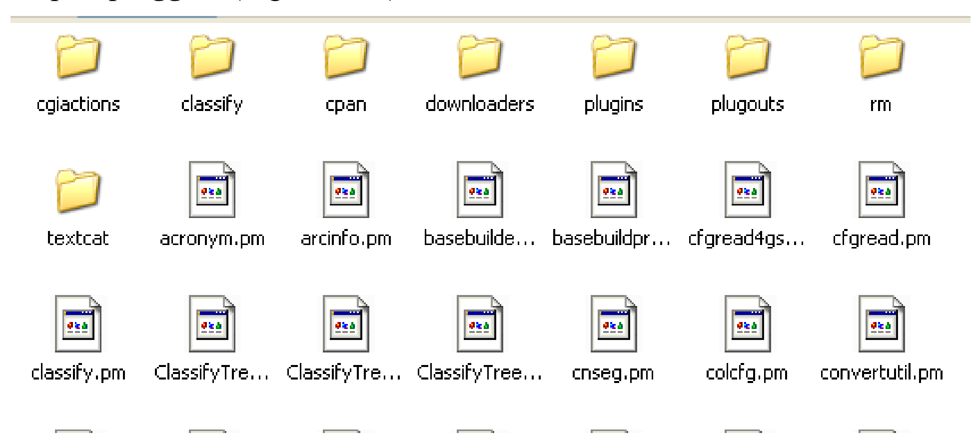


Figura 7.23. Contenido de la carpeta perlib.

Seria conveniente apuntar en este momento que existen dos tipos de pluggins: de metadatos y pluggins de documentos y no se produce ninguna distinción entre ellos dentro del sistema.

Todos los pluggins procesan files, y algunos aqrupan varios fles en un documento. Otros introducen un file en varios documentos

Los Plugins que procesan files XML por ejemplo son inherentes a XMLPlug. (Ellos impiden el sobreescribimiento del mismo código.)

Veamos el contenido de la carpeta pluggins (Figura 7.24).

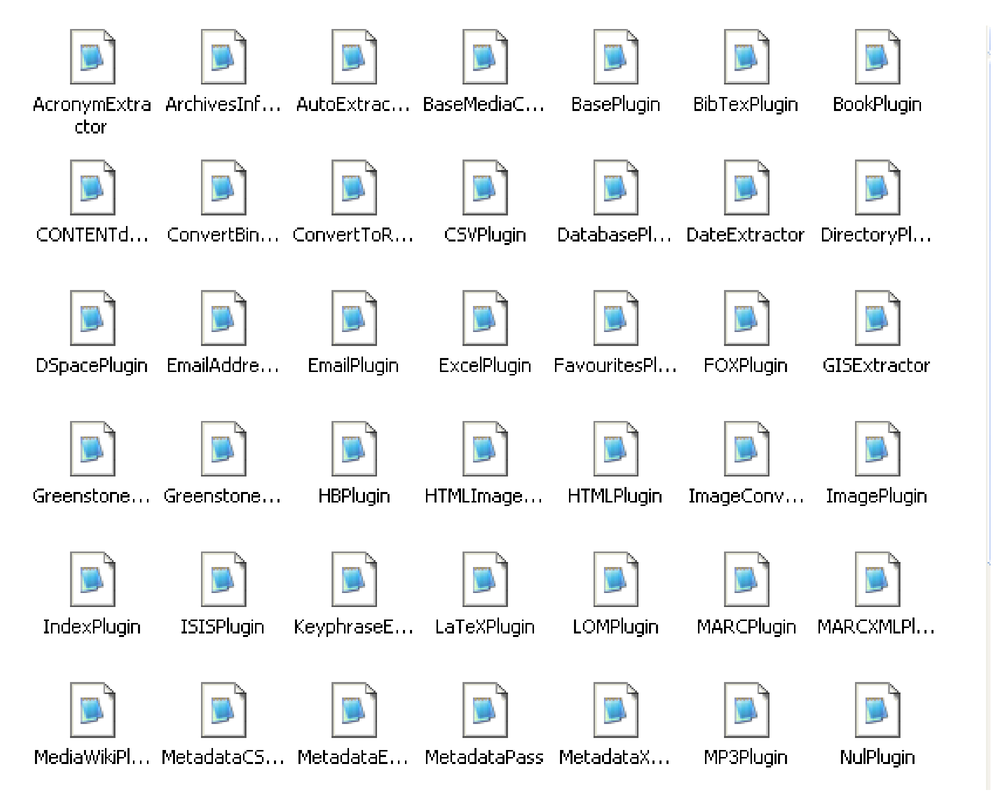


Figura 7.24. Contenido de la carpeta pluggins.

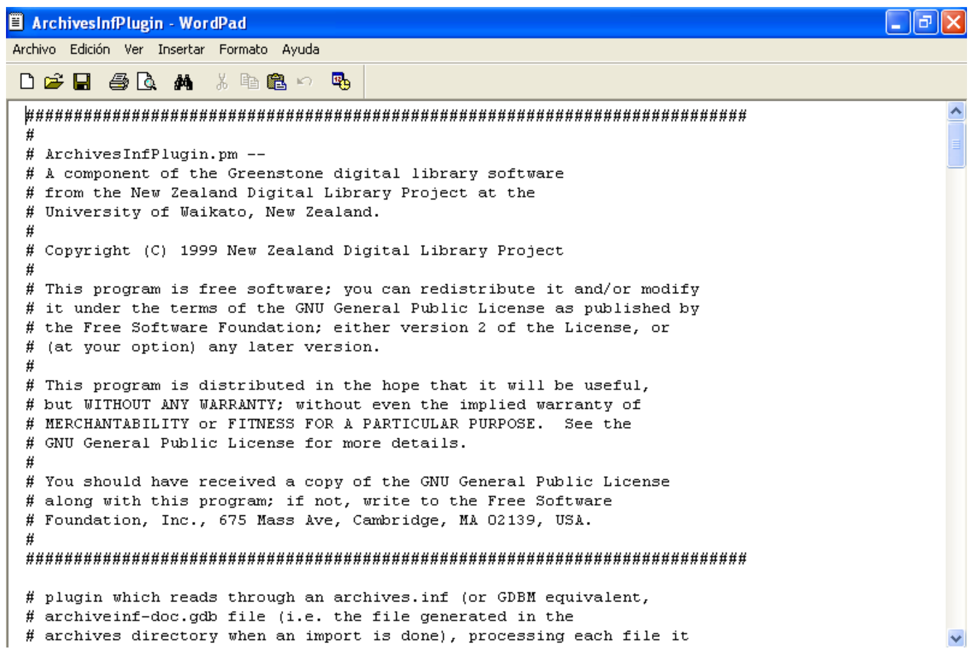
Plug-in es un programa que extiende las capacidades del navegador de un modo especifico, dado por ejemplo la capacidad de mostrar vídeo, audio, ficheros de un determinado formato (ficheros PDF, presentaciones de ASAP, fichero VRML, etc …). No existe actualmente un conjunto estándar de plug-in´s para cada tipo de ficheros, sino que existen diversas aplicaciones, realizadas por diversos fabricantes, y no todas de libre distribución. Se puede asegurar que todas las aplicaciones serán compatibles y si por ejemplo se referencia un fichero de sonido en formato .wav en su página, este podrá ser oído por todos aquellos que tengan un plug-in para este tipo de ficheros. Por tanto se recomienda un uso moderado de los plug-in´s utilizando ficheros que sean de formatos comunes y un uso complementario en una página, ya que muchos usuarios no podrán verlo (4).

Los conectores, o pluggins, analizan los documentos importados y extraen los metadatos en Greenstone . El conector HTML, por ejemplo, convierte las páginas HTML en el Formato de Archivo Greenstone y extrae los metadatos que están explícitos en el formato del documento, como los títulos entre etiquetas <title> </title>.

Los conectores se escriben en lenguaje Perl. Todos proceden de un conector básico denominado BasPlug, que realiza todas las operaciones necesarias como crear un nuevo documento en el Formato de Archivo Greenstone con que trabajar, asignar un identificador de objeto (OID) y manejar las secciones de un documento. Los conectores se guardan en el directorio perllib/plugins (5).

Veamos continuación las instrucciones que contiene el ArchivesInPluggin, abierto con el programa Wordpad (Figura 7.26).

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0.




```
###########################################################################
#
# ArchivesInfPlugin.pm --
# A component of the Greenstone digital library software
# from the New Zealand Digital Library Project at the
# University of Waikato, New Zealand.
#
# Copyright (C) 1999 New Zealand Digital Library Project
#
# This program is free software; you can redistribute it and/or modify
# it under the terms of the GNU General Public License as published by
# the Free Software Foundation; either version 2 of the License, or
# (at your option) any later version.
#
# This program is distributed in the hope that it will be useful,
# but WITHOUT ANY WARRANTY; without even the implied warranty of
# MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the
# GNU General Public License for more details.
#
# You should have received a copy of the GNU General Public License
# along with this program; if not, write to the Free Software
# Foundation, Inc., 675 Mass Ave, Cambridge, MA 02139, USA.
#
###########################################################################

# plugin which reads through an archives.inf (or GDBM equivalent,
# archiveinf-doc.gdb file (i.e. the file generated in the
# archives directory when an import is done), processing each file it
```

Figura 7.26. Contenido de ArchivesInPluggin, abierto con el programa Wordpad.

Veamos ahora el contenido de la carpeta Web de Greenstone (Figura 7.27).

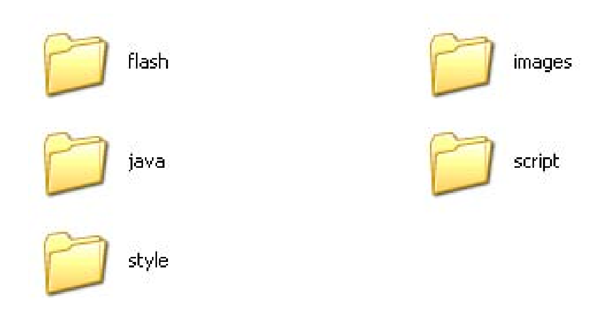


Figura 7.27. Contenido de la carpeta Web de Greenstone.

Destaquemos en esta carpeta la presencia de la sub carpeta style donde se alojan las hojas de estilo.

7.2.3- ¿Que son las hojas de Estilo y para que sirven?

CSS son las siglas de Cascading Style Sheets - Hojas de Estilo en Cascada - que es un lenguaje que describe la presentación de los documentos estructurados en hojas de estilo para diferentes métodos de interpretación, es decir, describe como se va a mostrar un documento en pantalla, por impresora, por voz (cuando la información es pronunciada a través de un dispositivo de lectura) o en dispositivos táctiles basados en Braille. CSS es una especificación desarrollada por el W3C (*World Wide Web Consortium*) para permitir la separación de los contenidos de los documentos escritos en HTML, XML, XHTML, SVG, o XUL de la presentación del documento con las hojas de estilo, incluyendo elementos tales como los colores, fondos, márgenes, bordes, tipos de letra ..., modificando la apariencia de una página Web de una forma más sencilla, permitiendo a los desarrolladores controlar el estilo y formato de sus documentos (6).

7.3- Sistema de ejecución de Greenstone. GSDLHOME. Cadenas de Formato

La interfaz gráfica de Greenstone es una herramienta de asistencia a los que elaboran las colecciones con este software. Se instala automáticamente con todas las versiones de Greenstone , Se coloca en el subdirectorio *gli* del nivel superior del Directorio Greenstone (8).

7.3.1- GSDLHOME

Los elementos definidos en el directorio *GSDLHOME/lib* son objetos Greenstone de nivel inferior, creados por encima de STL (o Biblioteca Standard de Plantillas, puede consultar para ampliar esta idea, el capitulo Términos Conceptuales y Referenciales) y que se utilizan en todo el código fuente. Así por ejemplo, el código fuente del sistema de ejecución se encuentra en *GSDLHOME/src.*

Los archivos de encabezados situados en el Directorio *GSDLHOME/lib* comprenden una mezcla de funciones y objetos que contribuyen útilmente al sistema de ejecución de Greenstone (2).

Greenstone funciona bajo Windows a partir de su versión 3.1, lo cual impone desafortunadamente un máximo de ocho caracteres en los nombres de archivos y directorios. Esto explica la utilización de abreviaciones crípticas como *recpt* y *colservr*. Los demás subdirectorios incluyen programas independientes que se utilizan principalmente para reforzar el proceso de creación (2).

Recordemos que hay dos archivos de configuración que controlan el funcionamiento de Greenstone: el archivo de configuración de sitio gsdlsite.cfg y el archivo de configuración principal *main.cfg.* (8).

El archivo *gsdlsite.cfg* se utiliza para configurar el programa Greenstone para el sitio en que está instalado. Está concebido para registrar las opciones de configuración propias de un sitio determinado como, por ejemplo, el nombre del directorio en que se guarda el programa Greenstone, la dirección HTTP del sistema Greenstone, y si se está utilizando o no la función fastcgi. La utilización, de la función fastcgi, permite que el programa library pueda permanecer en memoria después de la primera ejecución y recibir las subsiguientes series de argumentos CGI, con lo que se evita repetir la inicialización y su función se asemeja más a la de un servidor convencional. La utilización de Fast-CGI es una opción de Greenstone y se activa volviendo a compilar el código fuente con las bibliotecas adecuadas (2). En la Guía de Instalación de la Biblioteca Digital Greenstone se describen los elementos que contiene este archivo (8). Incluye la dirección de correo electrónico del encargado del mantenimiento del sistema, indica si las páginas de estado y del Colector están activadas o no, si se llevan o no los registros de las actividades de los usuarios y si se utilizan o no "*cookies*" para identificar a los usuarios.

7.3.2- Líneas del archivo gsdlsite.cfg

El archivo de configuración del sitio gsdlsite.cfg, que se encuentra en el mismo directorio que el programa library, establece variables utilizadas por el programa de biblioteca y el servidor Web durante la ejecución. Les recomendamos revisar las líneas de este archivo que podrá encontrar en un cuadro en la Sección 5 de la Guía de Instalación de la Biblioteca Digital Greenstone (8).

7.3.3- Archivo de configuración principal main.cfg

El archivo de configuración principal *main.cfg* sirve para configurar el recepcionista, esto es, la parte de Greenstone que procesa las consultas y muestra las páginas. Se puede controlar todo, desde las lenguas que la interfaz puede utilizar hasta los registros que se llevarán. El archivo de configuración main.cfg contiene dos tipos de entradas que inciden en la gestión de las diferentes

lenguas. Éstas determinan qué idiomas y qué códigos estarán disponibles en la página de preferencias (6).

El conjunto de la interfaz de usuario de Greenstone se controla mediante macros que se encuentran listados en el directorio *GSDLHOME/macros*. Estas macros están escritas en un lenguaje especialmente concebido para Greenstone y se utilizan en el momento de la ejecución para generar páginas Web. La traducción del lenguaje macro al formato HTML es la última etapa para mostrar una página. Así pues, las modificaciones de un archivo de macros afectan inmediatamente la presentación en pantalla y permiten experimentar rápida y fácilmente. Todos los archivos de macros que utiliza Greenstone se encuentran en *GSDLHOME/etc/main.cfg* y se cargan cada vez que se arranca el programa. La única excepción es cuando se utiliza la Biblioteca Local de Windows: en este caso es necesario reiniciar el proceso (6).

7.3.4- Cadenas de Formato

Las páginas Web que usted ve cuando utiliza Greenstone no han sido almacenadas previamente sino que son generadas "al instante" a medida que se las necesita. La apariencia de numerosos aspectos de las páginas se controla a través de las "cadenas de formato". Las cadenas de formato se encuentran en el archivo de configuración de la colección y se introducen con la palabra clave format seguida del nombre del elemento al que se aplica el formato (6).

Las cadenas de formato controlan dos tipos diferentes de elementos de página. El primero comprende los elementos de la página que muestran documentos o partes de documentos. El segundo incluye las listas generadas por los clasificadores y las búsquedas (6).

Todas las cadenas de formato se interpretan en el momento en que se visualizan las páginas. Es rápido y sencillo experimentar con las cadenas de formato ya que surten efecto en cuanto se guarda cualquier modificación en el archivo *collect.cfg.*

Ejemplo de cadena de formato:

"Expand Text|Expand contents|Detach|Highlight"

7.4- Interoperabilidad y Herramientas para la interoperabilidad en Bibliotecas Digitales. Estándares de metadatos

Destacan dos palabras en el encabezado de esta sección: interoperabilidad y estándares.

Desde un punto de vista informático, interoperabilidad se define como la habilidad que tiene un sistema o producto para trabajar con otros sistemas o productos sin un esfuerzo especial por parte del cliente. Este concepto tiene una importancia creciente a tenor de las colecciones digitales distribuidas que utilizan distintos esquemas de metadatos (9).

Un estándar se define como " Común y repetido uso de reglas, condiciones, lineamientos o características de los productos o procesos relacionados y producción de métodos, y manejos relacionados en la práctica sistemática (10)

Existen varios estándares y lineamientos que son importantes para así poder garantizar un valor presente y futuro de los documentos gestionados y su información, su expansión, integración e interconexión con o migración a otros sistemas (11).

La interoperabilidad entre distintos esquemas de metadatos puede realizarse de diversas formas, por ejemplo a través de el funcionamiento de un protocolo tipo OAI, obteniendo información mediante le pestaña descargar, lo que ampliaremos en el Capítulo 12, o bien a través del mapeo o establecimiento de correspondencias entre informaciones en diferentes formatos (por ej. MARC-

DC, FGDC-DC, que describiremos posteriormente) para la conversión de elementos de meta información que permita hacerlos compatibles (9)

Podemos hablar entonces del National Information Standards Organization Z39.50, Estándar DSpace, MARCXML y del Protocolo OAI-PMH, entonces como ejemplo de herramienta que posibilita la interoperabilidad de estos sistemas

7.4.1- National Information Standards Organization Z39.50

El protocolo Z39.50 adquiere parte del nombre de su creador, el comité número 39 de la American National Standards Institute (ANSI) y el .50, por ser el estándar número 50 por la National Information Standars Organization (NISO). Su nombre oficial es "Information Retrieval (Z39.50); Application Service Definition and Protocol Epecification for Open Systems Interconection", por simplicidad, en casi todos los textos se le denomina como Z39.50 (12).

Z39.50 es una norma para consultar catálogos de bibliotecas en Internet utilizando las mismas reglas para todos los catálogos, algo así como usar siempre la misma pantalla para consultar cualquier biblioteca (13).

Este estándar forma parte del sistema Greenstone y podemos constatarlo si nos dirigimos a la interfaz de trabajo del bibliotecario y abrimos la pestaña descargar (Figura 7.28).

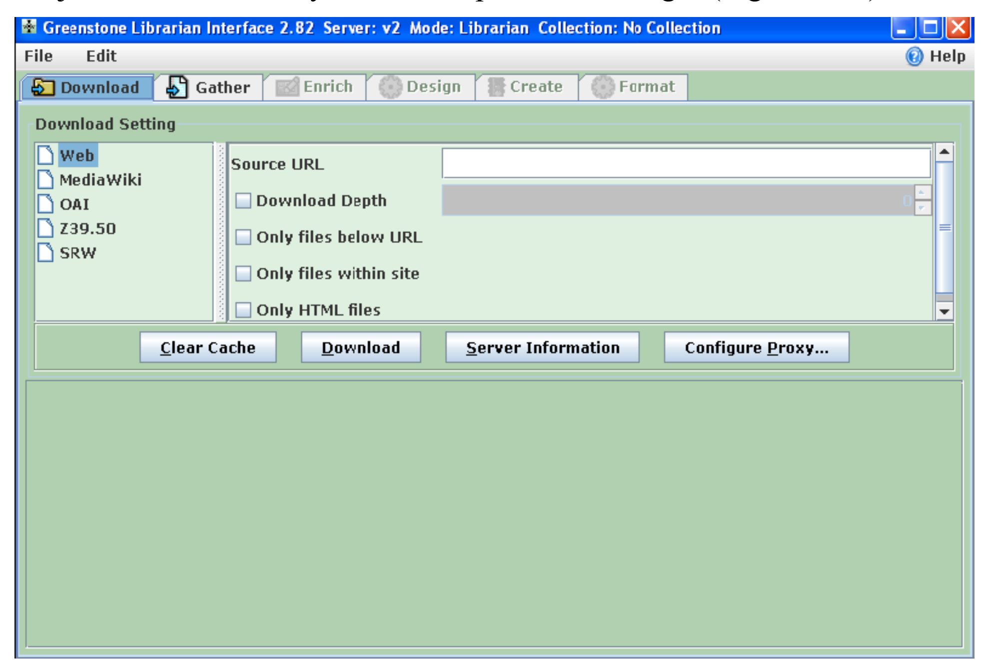


Figura 7.28. Interfaz de trabajo del bibliotecario, presencia del protocolo Z39.50 en Greenstone

Para consultar sistemas Z39.50 es necesario contar con un programa especial llamado cliente Z39.50 o navegador Z39.50, o bien a través de páginas o de portales en Internet. Es recomendable contar con un programa de consulta porque no todas las bibliotecas son accesibles a través de WWW (14).

Este estándar es mantenido por la Agencia de Mantenimiento Z39.50[2] que es administrada por la Biblioteca del Congreso de Estados Unidos (15)

[2] Disponible en: http://lcweb.loc.gov/z3950/agency

7.4.2- Estándar DSpace

DSpace es un software de código abierto diseñado por el *Massachusetts Institute of Technology* (MIT) y los laboratorios de HP para gestionar repositorios de ficheros (textuales, audio, vídeo, etc.), facilitando su depósito, organizándolos en comunidades, asignándoles metadatos y permitiendo su difusión a recolectores o agregadores. DSpace se una aplicación cliente/servidor que se gestiona vía Web, es decir, que la mayor parte de procesos pueden llevarse a cabo con un navegador estándar como Internet Explorer, Firefox u Opera. DSpace estructura los repositorios en comunidades y colecciones dónde las comunidades contienen sub-comunidades y/o colecciones y las colecciones contienen documentos.

Antes de poder instalarlo y habilitarlo es necesario instalar y configurar Sun Java, Apache Tomcat, Apache Ant, Apache Maven, PostgreSQL y Perl. Lamentablemente, Dspace no proporciona un paquete integrado o proceso de instalación automatizado que simplifique el proceso y permita instalar la aplicación y los software requeridos cómodamente (12).

Por defecto, DSpace está configurado con el esquema de metadatos Dublin Core, pero quizás, según los documentos que queramos depositar, será necesario disponer de otros esquemas para definirlos mejor (16).

7.4.3- Estándar MARCXML

Iniciativas de la *Network Development and MARC Standards Office* (Oficina de Desarrollo de Redes y Normas MARC) de la Biblioteca del Congreso de Estados Unidos para la conversión de registros de Catalogación MARC a marcado SGML (*Standard Generalized Markup Language*) o XML (*Extensible Markup Language*) y viceversa sin pérdida de datos (17).Ampliaremos sobre las características del formato MARC mas adelante.

7.4.4- OAI-PMH (*Open Archives Initiative - Protocol Meta-data Harvesting*)

OAI-PMH pretende:

- Desarrollar y promover estándares de interoperabilidad para la difusión de contenidos en Internet
- Mejorar el acceso a publicaciones electrónicas (eprints) y
- Objetivo más amplio: abrir el acceso a un rango de materiales digitales (18).

OAI-PMH es un protocolo de recopilación con un enfoque de interoperabilidad que proporciona una búsqueda distribuida que equivale a buscar y descubrir información y servicios remotos y recopilación, lo que equivale a que los datos/metadatos son transferidos desde la fuente remota al destino en el cual realizaran los servicio de búsqueda. OAI-PMH funciona sobre el protocolo HTTP. Las peticiones son operaciones HTTP GET o POST. Las respuestas son documentos XML válidos (18)

Este protocolo también forma parte del sistema Greenstone y podemos verificar su existencia en la interfaz de trabajo del bibliotecario en la pestaña descargar (Figura 7.29)

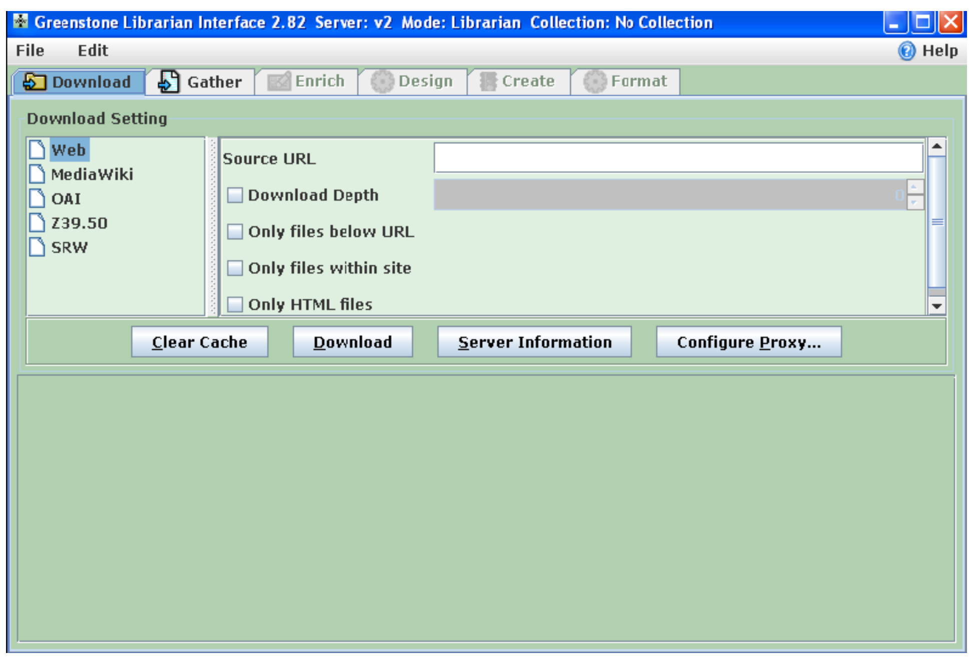


Figura 7.29. Interfaz de trabajo del bibliotecario, presencia del protocolo OAI-PMH en Greenstone

7.5- Características de la arquitectura del formato MARC en HTML en Greenstone

Henriette Avram, fue la creadora y quien implemento el formato MARC (*Machine Readable Cataloging*). Se le considera un vehículo para la comunicación de los datos bibliográficos y fue adoptado internacionalmente en 1973 (19).

Estas iniciativas han originado estructuras de datos en forma de DTD (*Document Type Definition o Definición de Tipo de Documento*) o Esquemas XML, herramientas de conversión, manuales y hojas de estilo. V. MARC en XML, en <http://www.loc.gov/marc/marcxmlspa.html>. Y asi por ejemplo les han sucedido (17):

- MARC SGML (1995)
- MARC XML DTD (2001)
- MARCXML Schema (2002) Las DTD son de gran tamaño, ya que cada elemento de datos MARC (campo o subcampo) es un elemento XML distinto. Para simplificar la estructura y mejorar la integración con aplicaciones XML orientadas a datos se ha desarrollado un Esquema XML (v. 1.0 2002-05-21, v. 1.1 2003-08-04). V. MARCXML : MARC 21 XML Schema <http://www.loc.gov/standards/marcxml//>

Les mostramos a continuación un esquema de esta estructura MARCXML[3] <MARC> (Figura 7.30).

[3] Disponible en: http://www.loc.gov/marc/

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0.

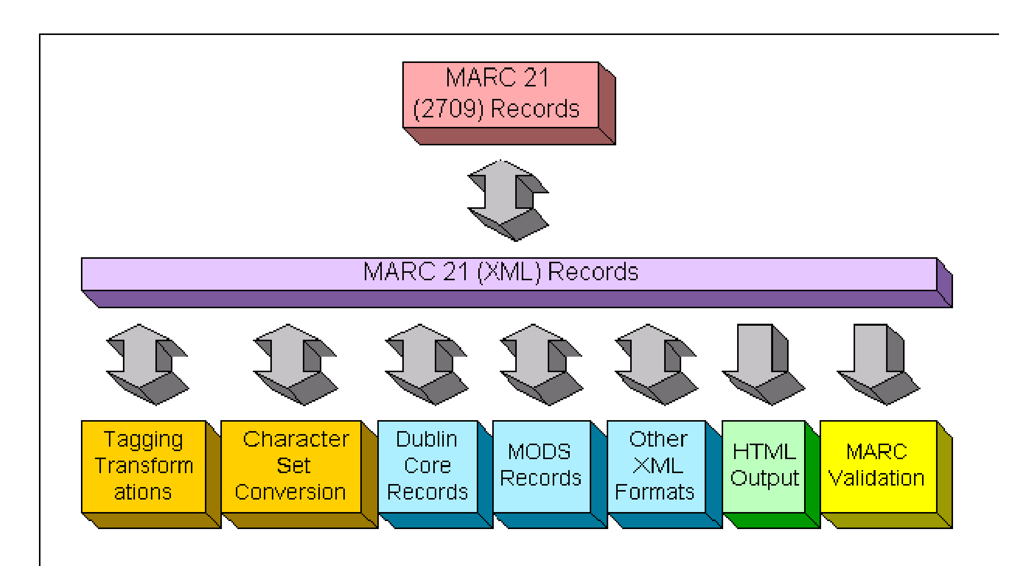


Figura 7.30. Esquema de la estructura MARCXML

Los principios de diseño que aplican son los siguientes (17):

1. Generalidad (independiente de aplicaciones específicas basadas en MARC o SGML),
2. Reversibilidad (la vuelta al formato original después de una conversión devuelve la semántica intacta),
3. Flexibilidad (son las aplicaciones las que pueden introducir restricciones e interdependencias entre elementos),
4. Amigabilidad (para usuarios bibliotecarios; así se han introducido elementos jerárquicos que explicitan asociaciones implícitas en el formato MARC; por ejemplo, los campos 2XX se agrupan en el elemento <mrcb-title-and-title-related>),
5. Relación con TEI (Text Encoding Initiative, en lo que se refiere a la cabecera TEI <teiheader>, haciendo complementarios ambos estándares)

Para una ampliación sobre el tema los invitamos a consultar el libro en Internet del Profesor J. Tomás Nogales Flores (17)

7.6- Consideraciones Finales

Estimamos que los contenidos de este capitulo representan las mejores piezas para la comprensión a fondo de las características de Greenstone, pues familiarizarse con la estructura interna del programa y los otros elementos que hemos destacado aquí, puede permitir una comprensión mas amplia de las singularidades de este sistema y provocar mayor confianza en los usuarios dispuestos a implementar este sistema como herramienta de trabajo.

7.7- Referencias Bibliográficas

(1) Wurman R 1975. Citado por Martín F.J. Hassan Y. Qué es la Arquitectura de la Información. No Solo Usabilidad [Seriada en Internet] no. 2, 2003/02/16. Disponible en: <http://www.nosolousabilidad.com/hassan >

(2) Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3. El Sistema de Ejecución de Greenstone [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

(3) Greenstone .org. Biblioteca Digital Greenstone. Guía del Programador .Configuración del sitio Greenstone. [Sitio Web en Internet] 2004. Disponible en:

<http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm>

(4) UCA.es. 7-CGI (Common Gateway Interface), ejecución de binarios [Libro en Internet].en: Manual HTML. 2002.Disponible en: < http://www2.uca.es/manual-html/ >

(5) Masternagazine. Definición de Plug-in.[Sitio Web en Internet] 16 de Febrero de 2005. Disponible en:<http://www.mastermagazine.info>

(6) Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 2. Explotación Óptima de sus Documentos. [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

(7) W3C.Hojas de estilo o cascading style sheets [Sitio Web en Internet] 2005. Disponible en: <http://www.w3c.es/divulgacion/guiasbreves/HojasEstilo>

(8) Greenstone.org. Guía de Instalación Greenstone. [Sitio Web en Internet]. 2002. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone >

(9) Mendez E. Senso A.Introduccion a los metadatos. Estándares y aplicación Conceptos asociados al uso de meta información . Unidad de Autoformación.[Sitio Web en Internet] SEDIC 2004. Disponible en: <http:www.sedic.es>

(10) Library of Congress. Technical Reports and Standards [Website in Internet]. 2005 Available from: <http://www.loc.gov/global/stdLauncher.js>

(11) Miranda C. Sistemas de Gestión de Documentos bajo Licencia de Código Libre [Sitio Web en Internet] 08 Junio 2009. Disponible en:<http://opengraphprotocol.org/schema//xmlns:fb>

(12) Cabrera A.M. Coutín A. Las Bibliotecas Digitales. Parte I. Consideraciones teóricas. Acimed [Seriada en Internet] 2005;13 (3). Disponible en: <http://bvs.sld.cu/revistas/aci/vol13_3_05/aci05305.htm>

(13) Biblioteca Daniel Cosío Villegas. Bibliotecas en Z39.50 [Sitio Web en Internet]. 2000. Disponible en: <http://biblio.colmex.mx/portal/estela14.JPG>

(14) Library of Congress. Biblio Tech Review. Z39.50: Part 1- an overview 12 [Website in Internet]. April, 2001. Available from: <http://lcweb.loc.gov/z3950/>

(15) Moen W. The ANSI/NISO Z39.50 Protocol: Information Retrieval in the Information Infrastructure. Biblio Tech Review.Information Technology for Libraries [Serials in Internet] 12 April, 2001. Available from: <http://syr.edu/50.brochure.part01.html>

(16) Rodríguez-Gairín J.M, Sulé A. DSpace: un manual específico para gestores de la información y la documentación.[Libro en Internet].Junio de 2008. Disponible en: <http://www.ub.edu/bid/20rodri2.htm >

(17) Nogales J.T. MARC en XML [Sitio Web en Internet] 2004-04-28. Disponible en: <http://www.bib.uc3m.es/~nogales/cursos/marcxml.html=intro>

(18) Van de Sompel, H., Lagoze C. The Santa Fe convention of the Open Archives Initiative D-Lib [Serials in Internet] . 2000, Febrero, v. 6, n. 2.Available from:<http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html>

(19) ALA | American Libraries. MARC Creator Henriette Avram Dies Libraries [Website in Internet]. April 26, 2006. Available from: < http://www.ala.org/Template.cfm?Section=SiteMap" >

CAPÍTULO 8

8- Construcción de una colección HTML con Greenstone. Modificación desde Greenstone Librarian Interface (GLI)

Lic. Elvia Vázquez Velázquez (elvia@cecam.sld.cu) Licenciada en Letras.
Centro de Cibernética Aplicada a la Medicina (CECAM) I.S.C.M.H. Dirección:
Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354

Palabras clave: Bibliotecas Digitales, Desarrollo Sostenible, Economía Digital, Estrategias de Inclusión Digital. , Sociedad de la Información, Software Libre, UNESCO.

8.1- Introducción

El proceso de creación y configuración de colecciones con Greenstone resulta más rápido y práctico por medio de la interfaz GLI, (Greenstone Librarian Interfaz) o Interfaz del Bibliotecario de Greenstone, una aplicación en Java, de entorno gráfico. Con esta se pueden reunir conjuntos de documentos, importarlos, asignarles metadatos y construir con ellos una colección, sólo debe dar un clic en cada una de las opciones que brinda para que pueda desarrollarla

8.2- ¿Donde localizar el GLI?

Para comenzar a trabajar, si trabaja con Windows como sistema operativo, va al botón Inicio, todos los Programas, y elige la versión que tiene instalada en su PC, en este caso Greenstone 2.82, selecciona Librarian Interface Greenstone (GLI).Tal como lo presenta la Figura 8.1



Figura 8.1 .Ruta donde se localiza GLI

8.3- Pegando figura

Inmediatamente aparece una consola con comandos MS-DOS (Véase Figura 8.2) y luego exhibe en pantalla la interfaz del bibliotecario con sus respectivos menús: *Archivo* y *Editar* como se aprecia en la Figura. 8.3

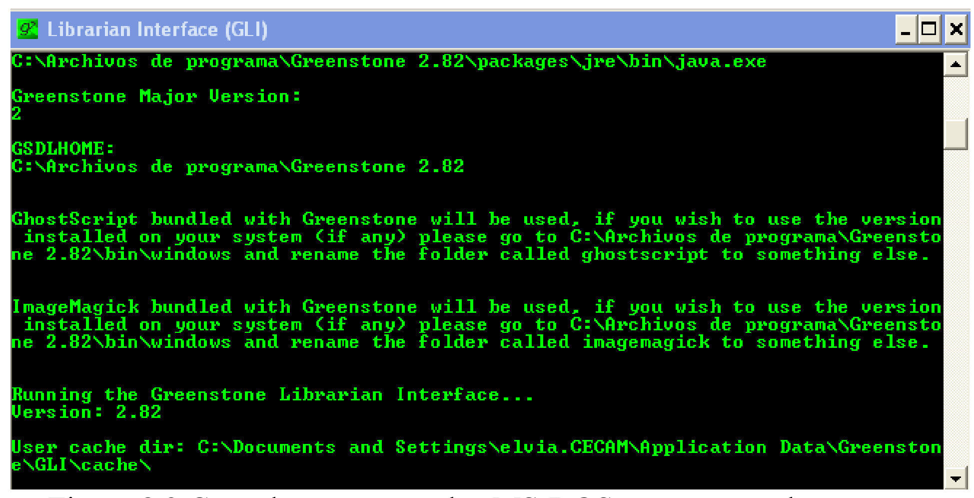


Figura 8.2 Consola con comandos MS-DOS para cargar el programa.

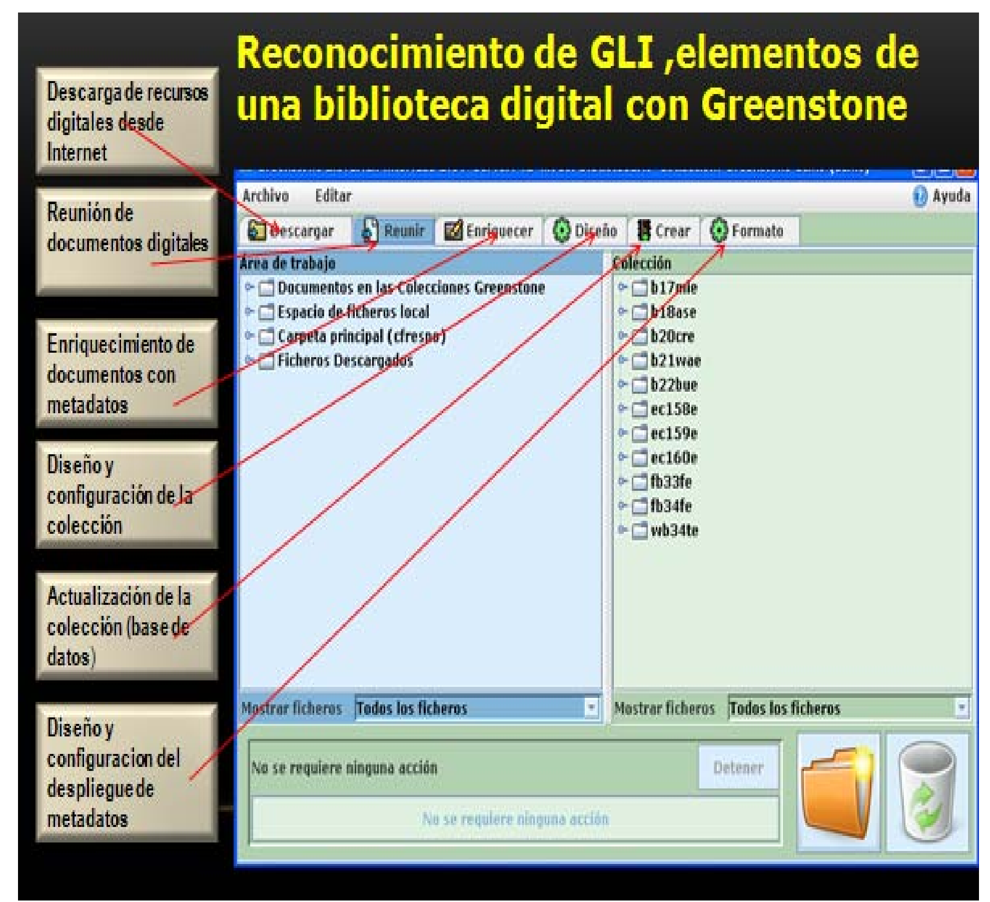


Figura 8.3 Pantalla que identifica GLI.

Para iniciar una colección debe de percatarse , dentro de la Opción ***Preferencias***, que aparece incluido en el menú ***Archivo***, el modo en que va a trabajar. Entre los cuatro posibles modos que posee esta interfaz seleccione el modo ***Avanzado***, pues accederá a todas las posibilidades u opciones que ofrece GLI. Fíjese en la Figura 8.4

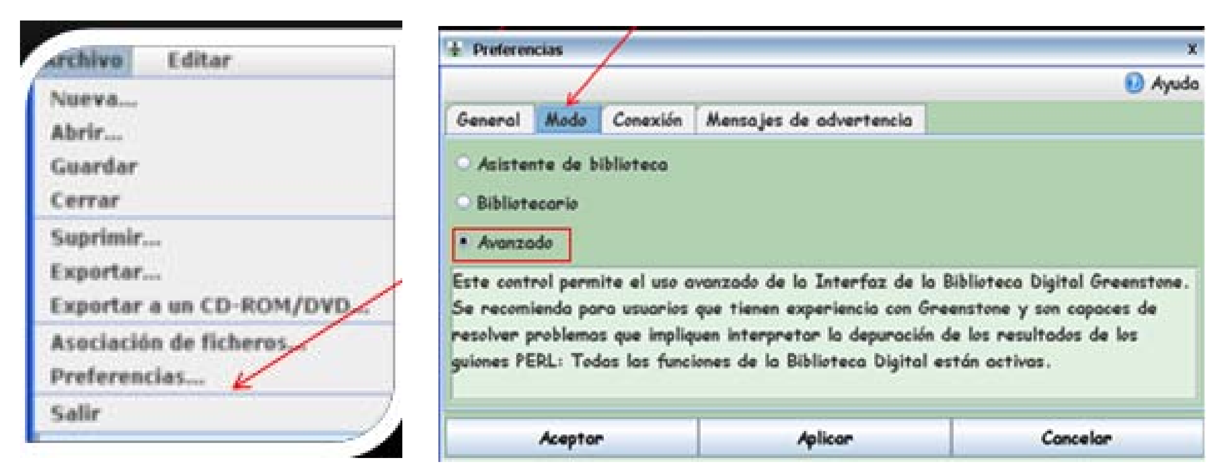


Figura 8.4 Opción Preferencias de Greenstone

Una vez elegido el modo ***Avanzado*** el siguiente paso es comenzar a crear una colección. Para ello de un clic en la opción ***Nueva*** dentro del menú ***Archivo*** como lo muestra la Fig. 5



Figura 8.5.Opción Nueva menú Archivo

Luego aparece el siguiente cuadro de diálogo de la Figura 8.6

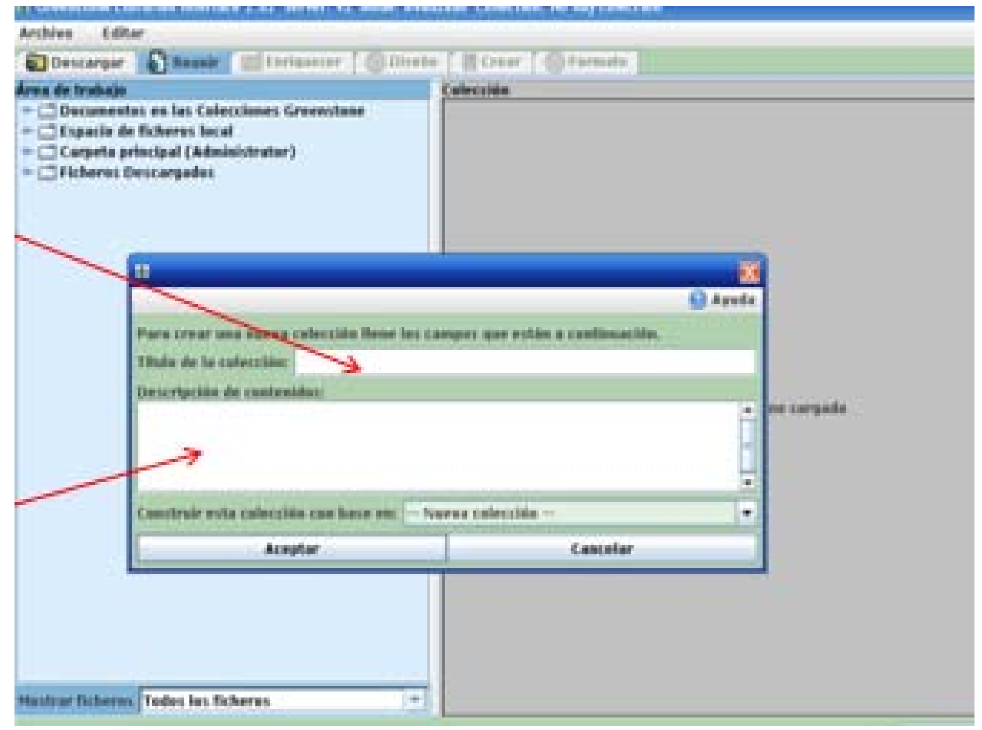


Figura 8.6 Cuadro de diálogo que indica creación de una nueva colección

Este le solicita al usuario que debe llenar la información general acerca de la colección en los campos correspondientes como queda indicado en la Figura 8.anterior (Véase Figura 8.6); el título de la colección debe ser una frase corta, en este caso la nombraremos "***Estrategia de Búsqueda***". Y en la casilla ***Descripción de contenido*** se describe detalladamente la información que identifica el tema o materia reunida en esta colección.

La pantalla presenta una imagen como la muestra la siguiente Figura 8.7.

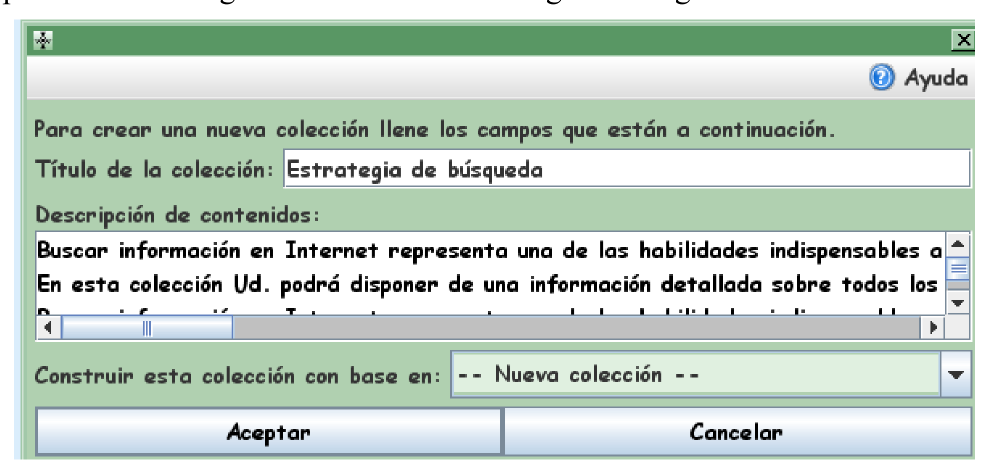


Figura 8.7 Definición de nueva colección

En el parámetro ***Construir esta colección con base en*** seleccione ***Nueva colección***.

Pulse sobre el botón Aceptar.

A partir de ahora todas las pestañas estarán activadas. Una vez realizados los procedimientos anteriores, surge un mensaje que le indica al usuario que comienza el proceso de creación de la colección. Ver Figura 8.8.

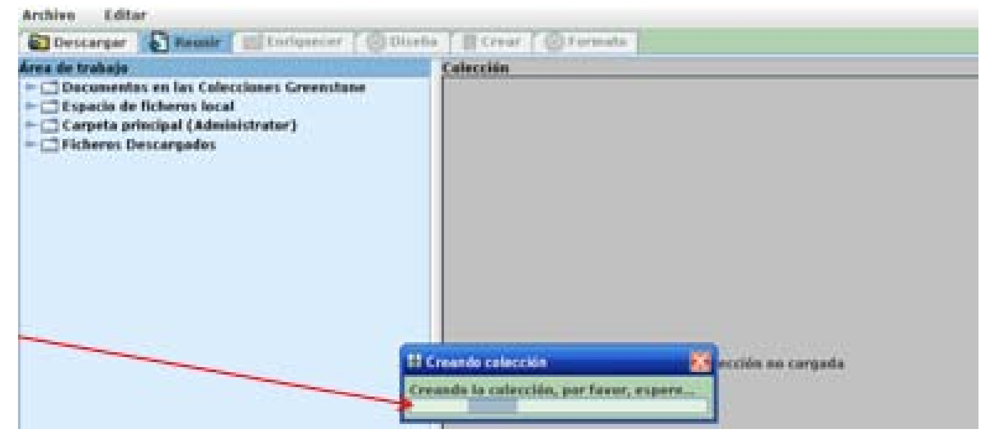


Figura 8.8 Proceso de carga de la colección

La pantalla del lado derecho de la interfaz cambia del color gris al verde, lo que significa que se ha activado el espacio de trabajo de Greenstone. Figura 8.9

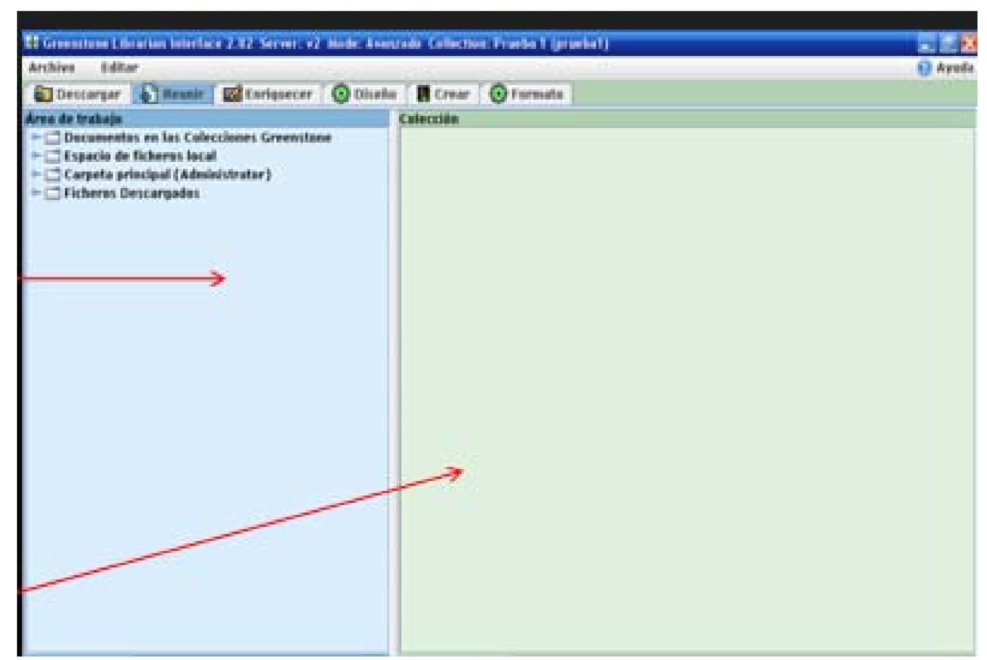


Figura 8.9 Espacio de trabajo de Greenstone color verde

De inmediato la pestaña ***Reunir*** se activa y toma el color azul; tiene dos áreas llamadas ***Área de trabajo*** y ***Colección*** que se ubican claramente debajo de las pestañas propias del software

8.4- Obtención de los documentos a incluir en la colección

El panel izquierdo o ***Área de trabajo*** permite explorar y localizar los documentos que se van a incluir en la colección y que se encuentran en su computadora. Una vez que haya localizado los documentos, lo que tiene que hacer es pulsar sobre estos, y arrastrarlos al panel derecho .Figura 8.9.

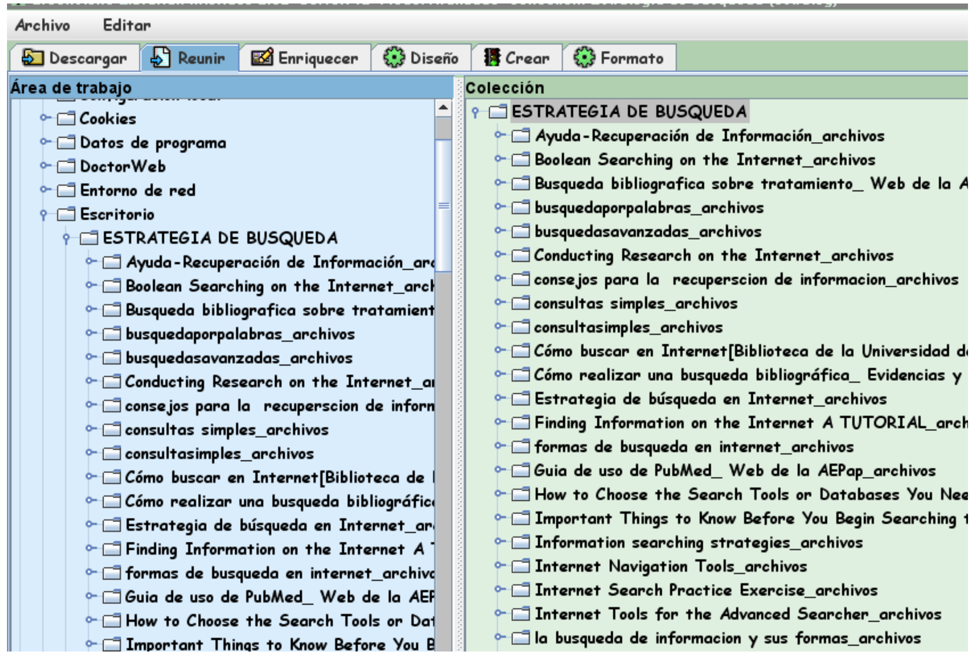


Figura 8.10.Selección de documentos existentes en el disco duro para incluirlos en colecciones

Para trasladar los documentos hacia el espacio de Greenstone (color verde) el usuario selecciona los documentos (izquierda, color azul), y con el mouse puede arrastrarlos y colocar los archivos en la parte derecha de la pantalla. Esta operación coloca los archivos convenientemente en la nueva colección, y en este momento los documentos heredan metadatos propios del sistema de forma automática. Estos se copian sin que para ello se afecte el sistema original de archivos existentes en su computadora.

Usted puede seleccionar varios archivos a la vez y surte el mismo efecto que si lo hace de forma individual. Cuando se seleccionen, arrastren y coloquen grandes conjuntos de archivos en la nueva colección, la acción de copiado puede llevar tiempo según la cantidad de documentos que usted importe. La interfaz muestra en la parte inferior, el grado de avance en que se está copiando el archivo y el porcentaje de archivos que se han procesado.

Para crear la colección de ejemplo se han seleccionado documentos en formato HTML.

8.5- Enriquecer los documentos con metadatos

El siguiente paso es enriquecer los documentos añadiéndoles metadatos para describir los documentos que hayamos añadido en la pestaña ***Reunir***. De esta manera se identifica cada recurso, y también sirve para recuperarlo posteriormente .

Este proceso se realiza en la pestaña ***Enriquecer*** (Ver Figura 8.11) despliega una nueva pantalla de información la cual exhibe los directorios y subdirectorios de la colección a la izquierda y a la derecha permite añadir los metadatos a los documentos individuales o a los grupos de documentos.

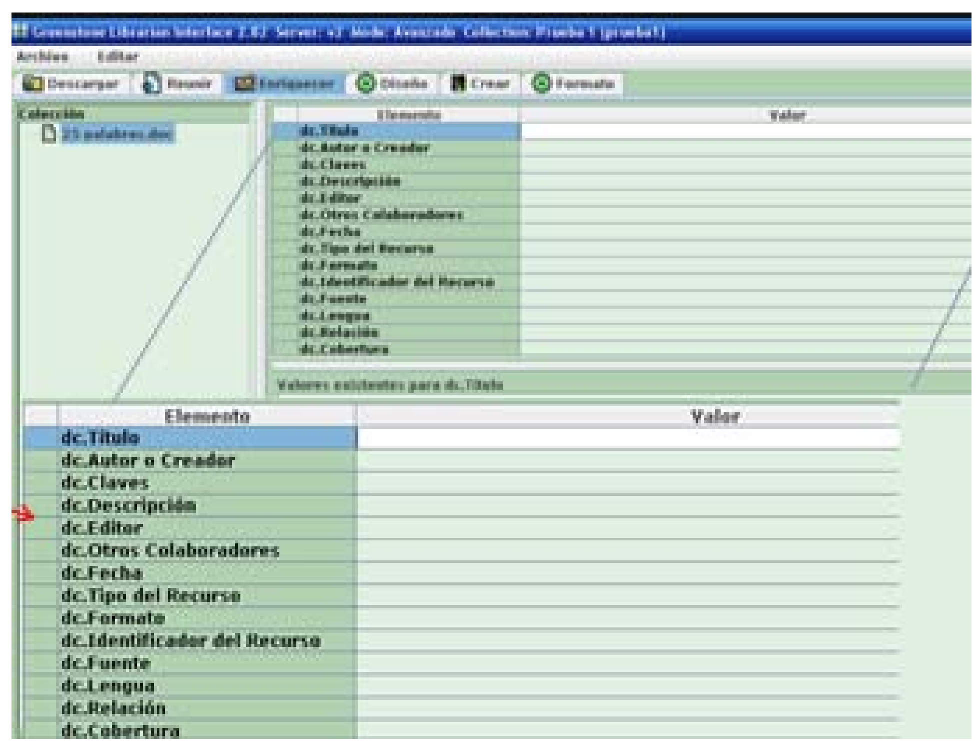


Figura 8.11 Pestaña “Enriquecer”, ingreso de metadatos

Ahora, el panel derecho de ***Enriquecer*** se ha activado, muestra el conjunto de metadatos seleccionado para los documentos. Los diferentes campos presentan una etiqueta que comienza por ***dc.*** (Correspondiente a Dublin Core), seguida del metadato que se trate (título, autor, descripción… etc.).

El conjunto de metadatos Dublin Core --Forma estándar de describir los metadatos (1), tales como el autor, el título, la fecha, palabras clave, etc. asociados a un documento o colección de documentos; es una norma que ha sido ampliamente aceptada como referencia de metadatos a escala mundial, puede ser incorporada a documentos XML (2) se compone de 15 elementos; pero no es necesario completar todos, permiten la descripción de cualquier recurso de información digital, atendiendo a aspectos de autoría y responsabilidad, descripción y contenido informativo-documental(3).Dada su flexibilidad, el bibliotecario tiene la posibilidad de incluir sólo aquellos elementos que le sean significativos; si sitúa el cursor encima de cada etiqueta de la columna ***Elemento*** (derecha) informa una ayuda sobre el contenido y la finalidad de cada metadato.

Durante el proceso de copiado arrastra consigo metadatos de forma automática. Si un documento es parte de una colección Greenstone, los metadatos definidos previamente pasan a formar parte de la nueva colección. Es posible que esta nueva colección tenga un conjunto diferente de metadatos, únicamente se conservan los que pertenecen al conjunto de la nueva colección. Con esta pestaña podemos asignar metadatos a un documento individual, o a una carpeta (estos son heredados por todos los documentos de la carpeta, incluye los que se encuentran en las carpetas anidadas).

Generalmente para los documentos electrónicos se utiliza el sistema de metadatos Dublin Core(4) sin embargo se contempla la posibilidad de utilizar otros sistemas.

El usuario también puede crear nuevos conjuntos de metadatos por medio de una pantalla que se activa a través de la pestaña ***Gestionar Conjunto de metadatos***. Véase Figura 8.12

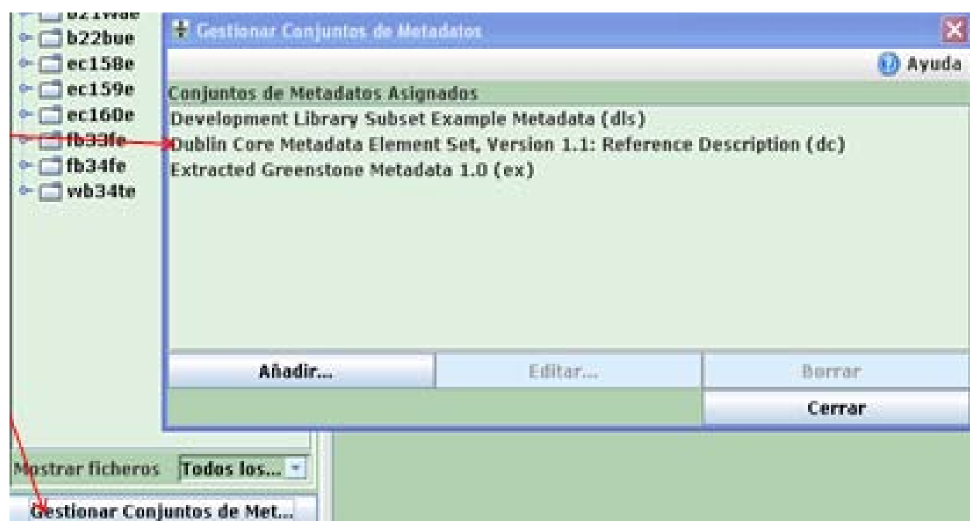


Figura 8.12.Modificar el conjunto de metadatos asignados

Puede revisar las características generales de estos sistemas en el capítulo sobre "Términos Conceptuales y Referenciales".

Diferentes conjuntos de metadatos se pueden asociar a la misma colección, ya que el sistema los mantiene separados (de tal manera que los documentos, por ejemplo, pueden tener un *Título* en Dublin Core y uno en DLS). Estos conjuntos se mantienen claramente diferenciados en la interfaz. Entre bastidores los conjuntos de metadatos se representan en XML. (4)

Para el ejemplo que proponemos hemos seleccionado llenar de forma manual los metadatos de ***Título, Autor o Creador, Palabras claves, Formato, Fecha.***

(Título): Normalmente, el título será el nombre por el que se conoce formalmente el recurso.

Autor o creador: Pueden ser, una persona, una organización, o un servicio.

Palabras clave: Generalmente, la materia se expresará como palabras clave, frases clave o códigos de clasificación que describan un tema/asunto del recurso.

Formato: El formato puede usarse para identificar el software, u otros equipamientos necesarios para visualizar /presentar u operar el recurso.

Fecha: La fecha se asociará con la creación o la disponibilidad [publicación en Red] de un recurso. (3)

La forma correcta de asignar metadatos es crear una entrada independiente para cada uno de ellos. Está dado para metadatos que ofrecen valores múltiples como: autor, editor, palabras clave, otros colaboradores, ¿Cómo obtener una entrada independiente? se escribe la información correspondiente y se pulsa "Enter", a continuación la siguiente entrada que queremos brindar al programa en ese mismo metadato.

8.6- Diseño y configuración de la colección

El usuario revisa la información acerca de la colección ingresada al momento de crearla y tiene en cuenta el diseño y configuración de esta en el que expone su gusto, y conocimiento de diseño, para ellos debe activar la pestaña Diseño.

Esta pestaña ***Diseño*** le permite especificar la estructura, organización y presentación de la colección que se crea. Este paso incluye una serie de pantallas separadas, cada una de las cuales trata un aspecto relativo al diseño de la colección y sirve, en efecto, como un equivalente gráfico al proceso

normal de editar manualmente el archivo de configuración. (5).

En la Figura 8.13, a la izquierda, aparece una lista de las diferentes interfaces que el usuario puede configurar para lograr el diseño final ellos son: ***Plugins de documento***, ***Índices de búsqueda***, ***Particionar índices***, ***Clasificadores de exploración***.

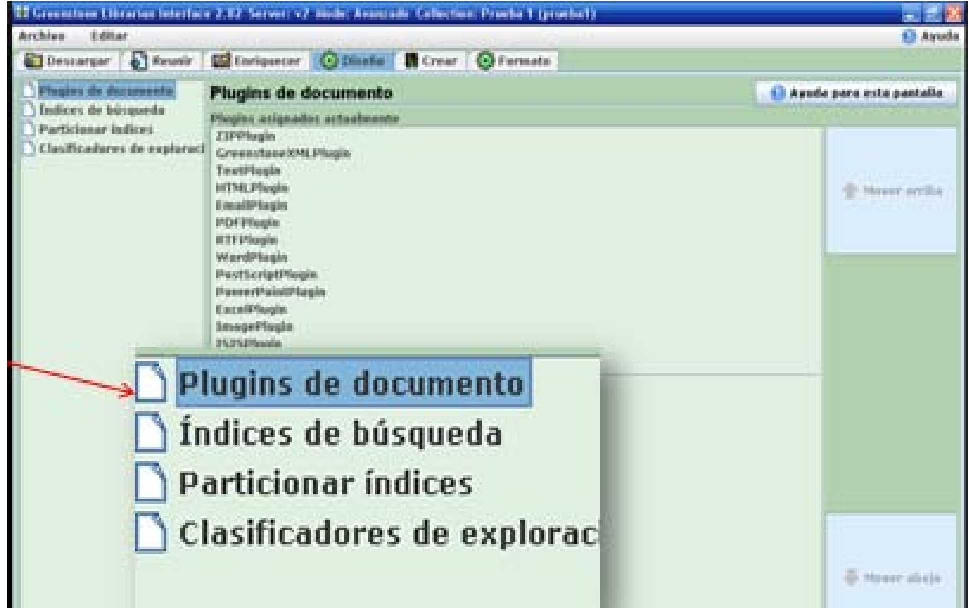


Figura 8.13. Interfaz de pestaña Diseño

La apariencia y sus funciones varían entre ellas, por ejemplo, al hacer clic en el botón ***Plugins** de documento*, que aparece representado en la Figura 8.14 , permite añadir, eliminar o configurar los plugins o conectores que trae por defecto el programa, así como cambiar el orden en el que se aplican a los documentos.

Los plugins o conectores (6) son pequeños programas que tratan documentos en texto sin formato, html, word, pdf, los mensajes usenet y de correo electrónico, etc. y los convierte en un formato de archivo Greenstone, estilo XML. Son los plugins quienes reconocen el formato de los documentos originales para procesarlos de manera igual durante el progreso de diseño y construcción de la colección.

Los conectores o plugins se escriben en lenguaje Perl (Practical Extraction and Report Languaje) (7), un lenguaje pensado para la manipulación de cadenas de caracteres, archivos y procesos. Debido a sus características Perl es muy potente en la creación de pequeños programas que pueden ser usados como filtros para obtener información de ficheros, realizar búsquedas, etc.

El programa Greenstone originalmente está dotado de varios plugins como lo presenta la Figura 8.14.



Figura 8.14.Plugins de documentos asignados por Greenstone

Seleccione el Plugins según sea el formato del documento original, para el caso que se ejemplifica escogemos el Plugins ***HtmlPlugin*** pues los documentos que se procesan son formato HTML. Para ello debe seleccionarlo en el extremo inferior derecho del listado que le ofrece esta interfaz .Ver Figura 8.16.

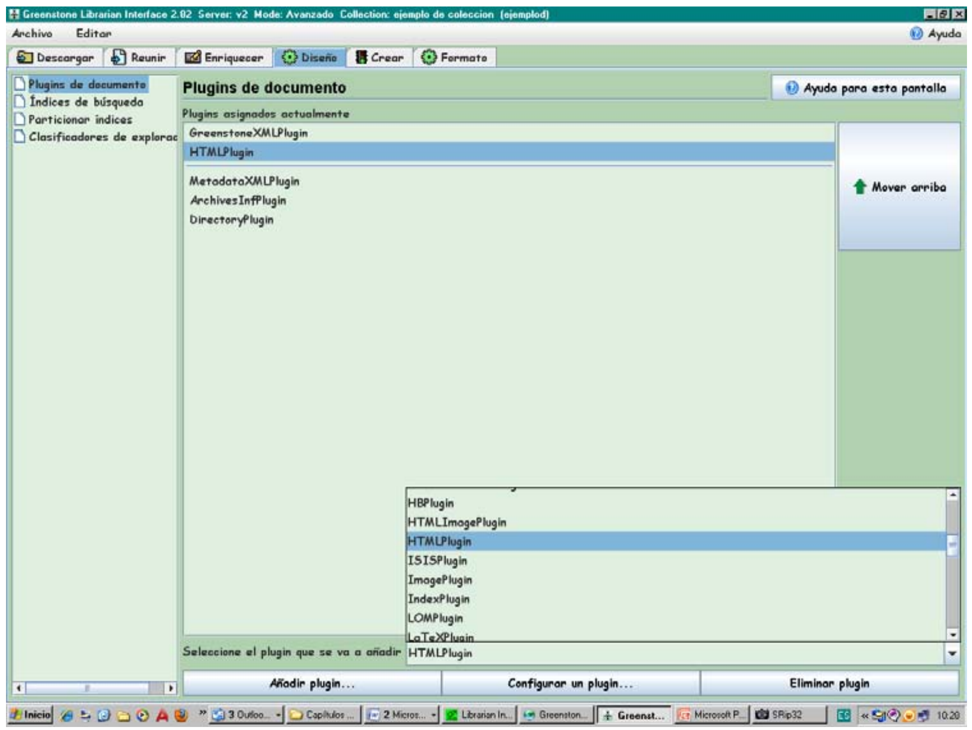


Figura 8.15. Opción Plugins de documentos

Puede eliminar los que no va a utilizar excepto *Greenstone XML Plugin* y los que aparecen bajo la línea divisoria de esta pantalla, es decir ***Metadata XML Plugin***, ***ArchivesInfPlugin***, ***DirectoryPlugin***; necesarios para que funcione Greenstone.

Igualmente es aconsejable añadir: TextPlugin y HTMLPlugin, pues ayudan considerablemente a procesar toda la información.

Cuando lo selecciona de la lista que aparece en el extremo inferior derecho al dar doble clic sobre él, aparece un cuadro de diálogo con varios argumentos que permiten especificar el plugin y refinar la configuración de este. Véase Figura 8.17

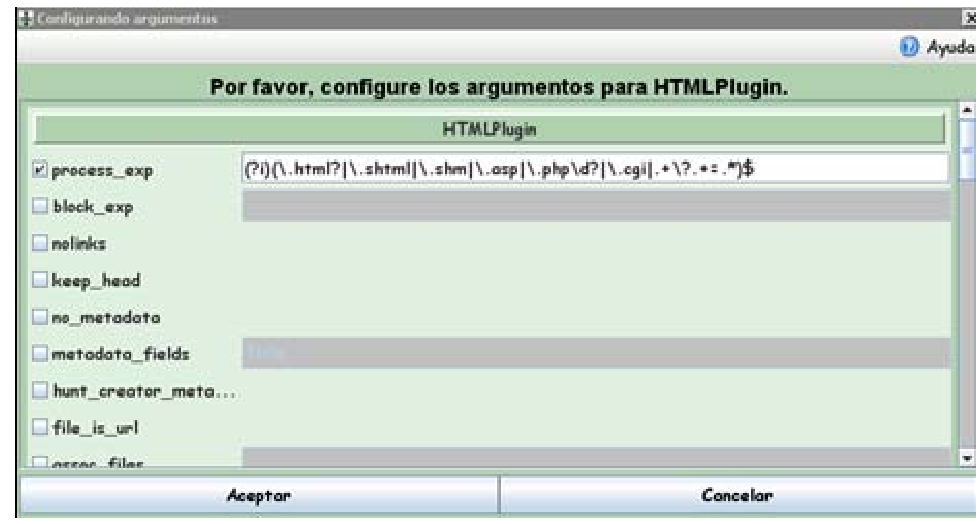


Figura 8.16.Cuadro de diálogo con argumentos para configurar Plugin

Para configurar los plugins debe activarse con un shake mark el argumento ***process_exp*** (que aparece señalado en la figura) resulta más útil la función del plugin a utilizar. Del resto usted selecciona según lo necesite para perfeccionar la configuración correspondiente. Cada uno ofrece opciones y argumentos diferentes que se pueden modificar para desarrollar mejor su colección. El conector HTML, por ejemplo, convierte las páginas HTML en el Formato de Archivo Greenstone y extrae los metadatos que están explícitos en el formato del documento, como los títulos entre etiquetas <title> </title>.

Cuando va a crear una colección con documentos HTML, como el modelo a tratar en este capítulo, si durante el proceso de importación de documentos originales, GLI advierte que existe algún documento con extensión o formato con el que Greenstone no es compatible aparece una pantalla que lo señala y sugiere utilizar el plugin ***UnknownPlugins*** tal y como lo presenta la Figura 8.18

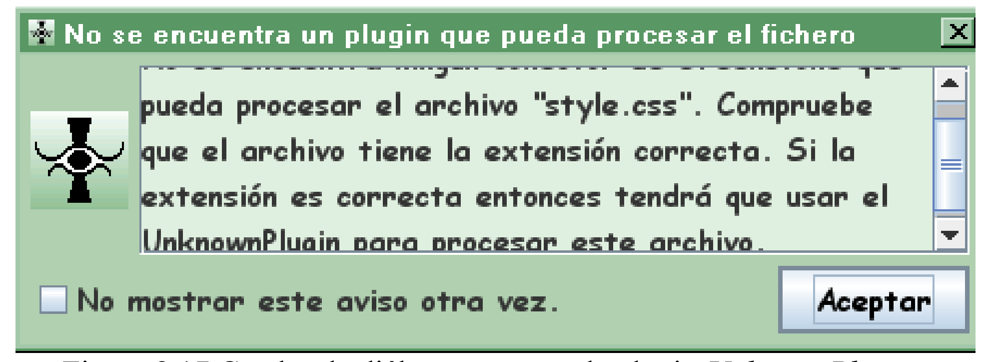


Figura 8.17.Cuadro de diálogo para uso de plugin *UnknownPlugins*

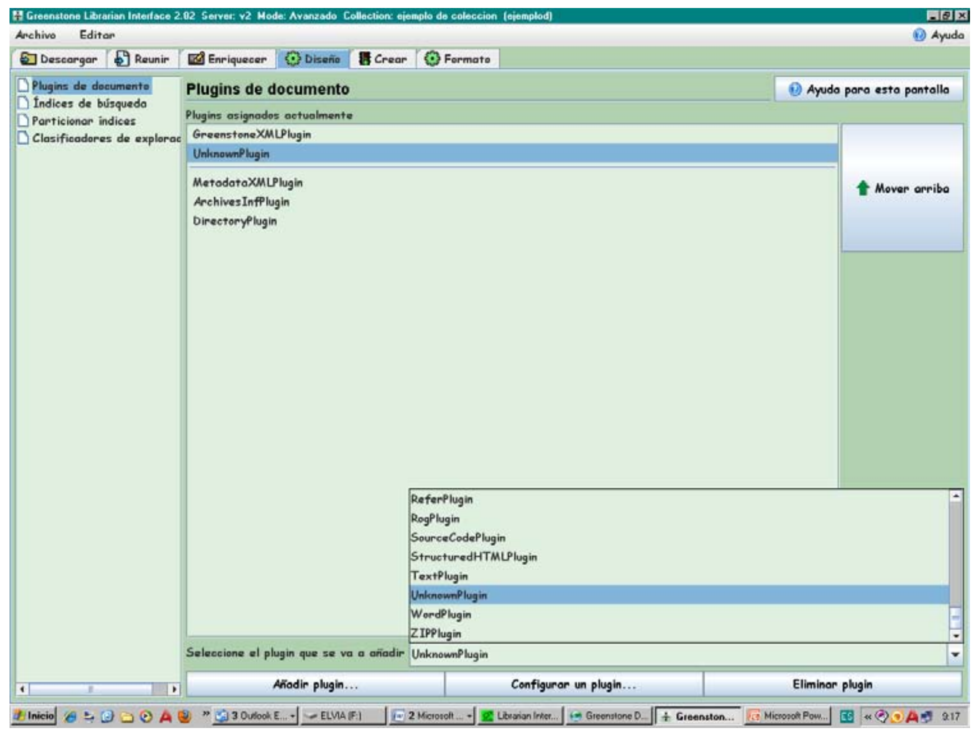


Figura 8.18.Selección de plugin *UnknownPlugins*

En el ejemplo informa que existen documentos con extensión ***.css .jb*** y que Greenstone no los reconoce por lo tanto sugiere usar ***UnknownPlugins***

Para ello selecciónelo de la lista de plugins y configure con el nuevo tipo de extensión que se importó.

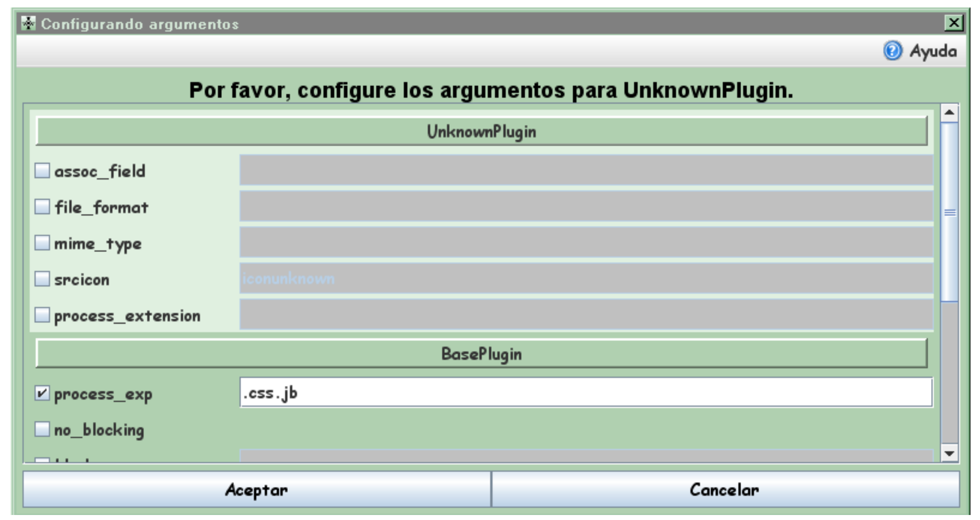


Figura 8.19.Configuracion argumentos UnknownPlugin

El uso de los Plugin puede resultar complicado, por eso hay que ensayar con todos los argumentos y conocer muy bien las posibilidades de cada uno de ellos.

8.7- Índices de Búsqueda o Consulta

Esta opción determina en qué parte de los documentos puede buscar el usuario posteriormente.

Preliminarmente Greenstone posee tres ***índices de consultas o búsqueda***: ***texto, título y fuente***

(Figura 8.20), estos pueden ser eliminados y cambiados con un clic en la pestaña ***Nuevo índice***. Estos índices se seleccionan según el elemento de metadato especificado.

Los índices se pueden construir basados en más de un elemento de metadato como por ejemplo el caso de Autor y Autores corporativos.

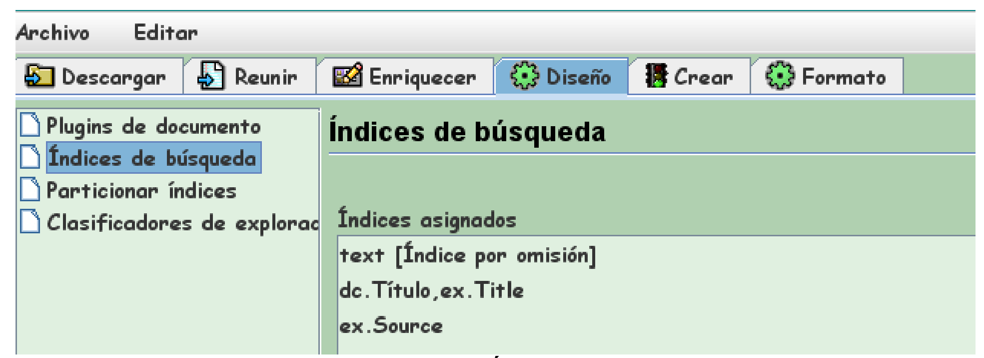


Figura 8.20. Opción Índices de Búsquedas

Para añadir un índice de búsqueda, selecciónelo de la lista ***Índice asignados*** (Ver Figura 8.21) y haga clic en ***Nuevo Índice***.

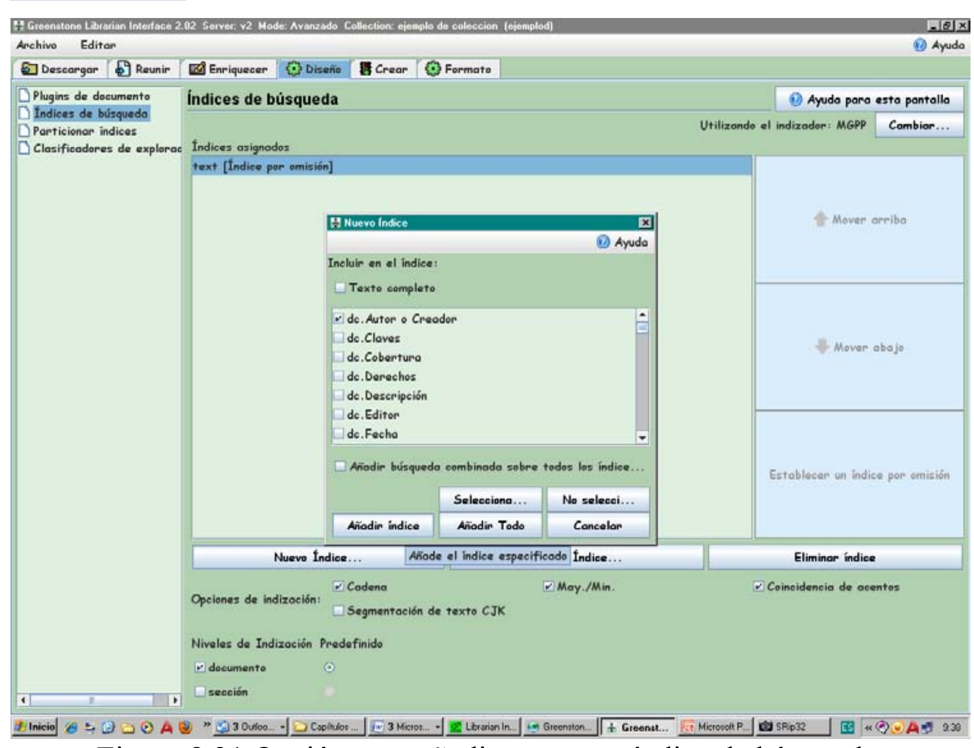


Figura 8.21.Opción para añadir un nuevo índice de búsqueda

Cada uno puede aparecer en la lista solamente una vez. Para quitar alguno selecciónelo de la lista de búsqueda asignados actualmente y haga clic en ***Eliminar índice de búsqueda***.

Para cambiar el orden de un índice de búsqueda, selecciónelo y haga clic en "***Mover arriba***" o "***Mover abajo***" que se encuentran del lado derecho de la pantalla. Podemos añadir búsquedas combinadas en todos los índices que buscará en todos los apartados del documento, aparecerá en el listado como ***allfields***.

Para el ejemplo se han modificado los índices originales y se añadieron índices de **título, autor**, **palabras clave** y **fecha** tal y como puede ver en la Figura 8.22 que a continuación se muestra.

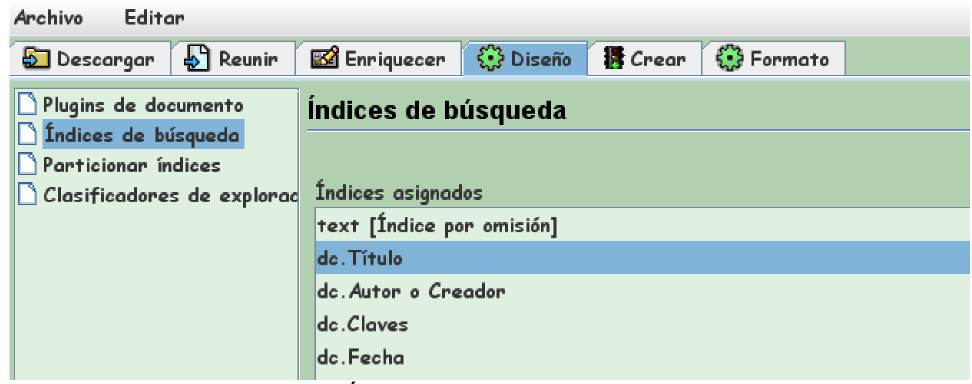


Figura 8.22. Índices de Búsqueda seleccionados

El orden que tengan los índices en la lista de ***Índices asignados*** durante el proceso de configuración de la colección, será el mismo orden en que aparecerán en el menú de la página de Búsqueda cuando usted despliegue en la casilla de ***Buscar en*** como se aprecia en la Figura 8.23.



Figura 8.23.Resultado del orden de índices en la lista de *Índices asignados*

8.8- Clasificadores de exploración

Corresponde entonces ahora activar la opción ***Clasificadores de exploración*** (Ver Figura 8.24) esta es la forma de asignar "clasificadores"estos sirven para crear los índices de consulta de la colección, los cuales se usan para navegar y buscar en ella.

Son las formas de mostrar los documentos de la colección al usuario, según los criterios de ordenación predefinidos. Los clasificadores son pequeños programas escritos en lenguaje Perl (recordar lo descrito sobre lo relacionado con los ***Plugins***) y se almacenan en el directorio perllib/classify. Ofrecen una ordenación y clasificación de todos los documentos. Se utilizan durante la creación de la colección.



Figura 8.24 Opción Clasificadores de exploración

Greenstone presenta dos clasificadores esencialmente, estos los puede anular para seleccionar nuevos del listado original.

Se despliegan varios clasificadores, se selecciona el que usted desee.

Para quitar uno de la lista, selecciónelo y haga clic en ***Eliminar el clasificador seleccionado***. Para añadir un clasificador nuevo, elíjalo por medio de la lista que ofrece la pantalla en el extremo inferior derecho y haga clic en ***Añadir el clasificador especificado*** presente en Figura 8.25.

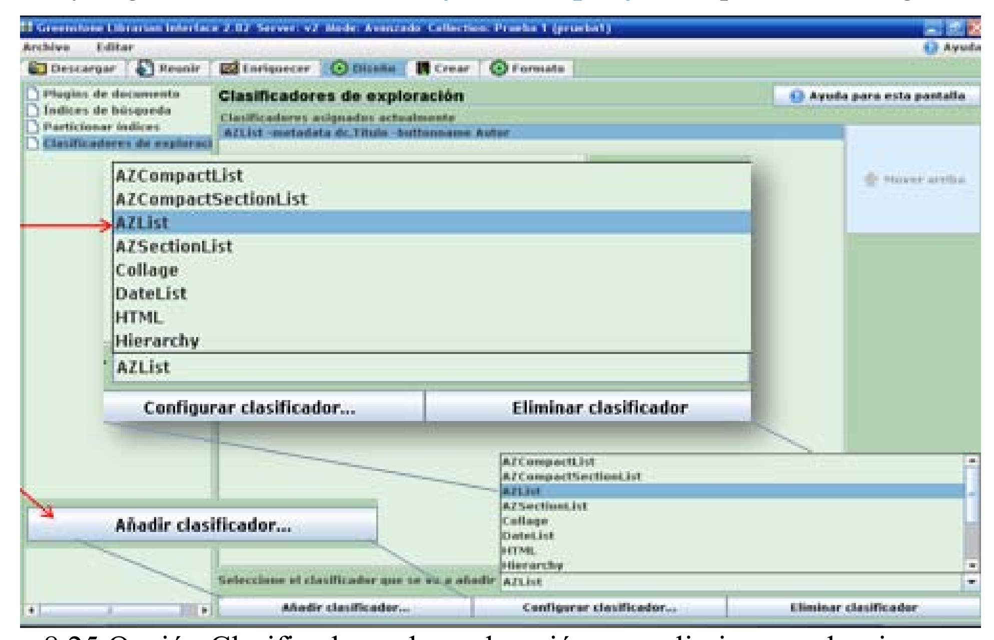


Figura 8.25 Opción Clasificadores de exploración para eliminar y seleccionar nuevos.

Con esto aparecerá una ventana titulada ***Argumentos de configuración***, las instrucciones para este cuadro de diálogo son las mismas que tienen los plugins. Una vez que haya configurado el nuevo clasificador, éste se añadirá al final de la lista de ***Clasificadores asignados* actualmente** tal y como lo presenta la Figura 8.26.

Para la colección de muestra se ha escogido el clasificador ***AZList –metadata Título***, (Véase fig.27) este crea una lista alfabética de títulos en el que reúne todos aquéllos que están dotados del campo de metadatos **Título** y los clasifica y divide por orden alfabético.

Ahora debe seleccionar la opción ***Configurar argumentos***, en la pantalla se despliegan una serie de opciones que permiten refinar los atributos del clasificador, se selecciona el metadato por el cual se basará el clasificador, según seleccione , en el caso de ejemplo se ha optado por el metadato Título. Figura 8.26

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0



Figura 8.26 Opción para configurar Clasificador de exploración *AZList*

Si no se especifica un argumento metadato, todos los documentos se incluyen en el clasificador en el orden en que aparecieron durante el proceso de creación; sin embargo con la elección del metadato adecuado y visto en la Figura 8.26, el resultado presenta una lista con los documentos ordnados alfabéticamente por título.

Algunos tipos de clasificadores:

- List: lista alfabética simple
- AZList: lista ordenada de la A a la Z
- AZCompactList: lista ordenada de la A a la Z, agrupa según número de documentos
- DateList: lista ordenada por fecha
- Hierachy: lista jerárquicas, por orden de metadatos
- …Section…: usados cuando se trata de documentos con secciones y subsecciones

Preferimos ahora el clasificador ***AZCompacList,*** (Figura 8.27) que sirve para los campos de ***Autor*** y ***Palabras clave***, este genera estanterías, y coloca una lista de todos los autores en el elemento de *metadato Autor*. Es decir agrupa los documentos según el autor dentro de cada estante.



Figura 8.27 Selección de clasificador AZCompacList

Otro clasificador que designamos es DateList, también visible en la Figura 8.anterior, genera una lista de selección de intervalos de fechas Figura 8.27.

Todos los clasificadores generan una estructura jerárquica que se utiliza para mostrar un índice de consulta.

8.9- Pestaña Formato

8.9.1- Opción General

En la pestaña ***Formato***, se encuentra la opción *General* (Figura 8.28) la cual demanda una sucesión de datos para completar la información del documento de la colección, como son: ***el correo electrónico del administrador de la colección, el título de la colección*** que se supone que ya usted lo escribió cuando comenzaba a crear la colección y ahora aparecen, ***los enlaces o URL*** para otorgarle un ícono o imagen que será el logo que identifique a esa colección y la opción para decidir si la colección será ***pública o no*** (Ver Figura 8.28).

Esta pantalla le permite a usted modificar el texto de la descripción de la colección que usted escribió cuando iniciaba el proceso de creación de la colección si fuera necesario efectuar algún cambio.

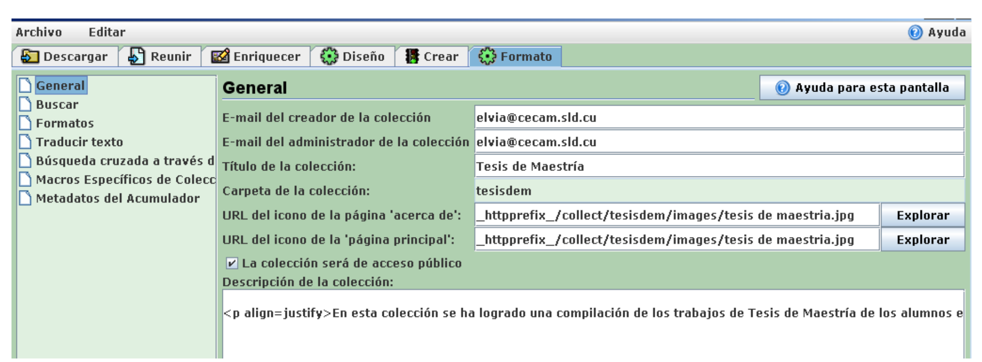


Figura 8.28 Pestaña Formato” ,Opción General

8.9.2- Opción Buscar

Esta opción está relacionada con los Índices de búsqueda de la pestaña Diseño.

En este debe cambiar los datos que aparecen en inglés y editar los nombres (en español) para los índices de búsqueda (Figura 8.29) así cuando el usuario realice una búsqueda de información puede optar por el más favorable para él. Estos son los que se visualizan posteriormente en la colección en la función ***Búsqueda***.

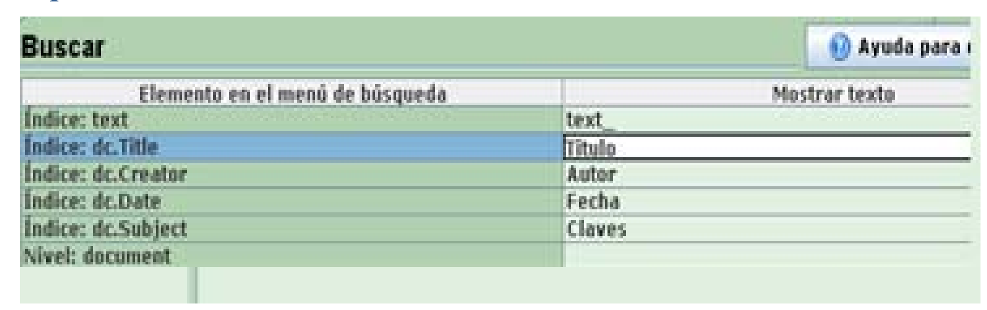


Figura 8.29 Edición de índices de búsqueda en español

La imagen en pantalla presenta el orden de los índices de búsqueda, en idioma español y son los que usted configuró previamente en la opción ***Índices de búsqueda*** .Para el usuario quedaría de la forma que lo representa la Figura 8.30 siguiente.



Figura 8.30 Visualización de índices de búsqueda en la colección

8.9.3- Opción Formato

En esta opción se despliega una pantalla con una cadena de caracteres en la que se combinan: funciones, componentes, variables con etiquetas HTML y etiquetas propias de Greenstone. Ver Figura 8.31

Existe un listado superior que recoge los distintos formatos que dispone la colección, ***Hlist***, ***VList***, y ***DocumentButtons***, determinan la presentación de los registros al navegar los Clasificadores de Exploración que se hayan escogido en la configuración de la colección. ***Document Heading***, y ***Documen***t ***Text*** son formatos que ofrecen el contenido completo de cada documento, según las opciones seleccionadas. Ver Figura 8.31

Formatos

```
VList "<td valign="top">[link][icon][/link]</td><td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[
HList "[link][highlight][ex.Title][/highlight][/link]"
DocumentHeading "{Or}{[parent(Top):Title],[Title],untitled}<br>"
DocumentText "[Text]"
DocumentButtons "Detach|Highlight"
SearchTypes "plain,form"
```

Figura 8.31.Opción Formato. Cadenas de Formato

También aparece un menú desplegable en la parte inferior:

Elegir función, que sirve para determinar las posibilidades del formato,

El Componente afectado determina el formato de la lista al que se refiere.

Cadena de formato HTML, brinda el código y los detalles de cada formato.

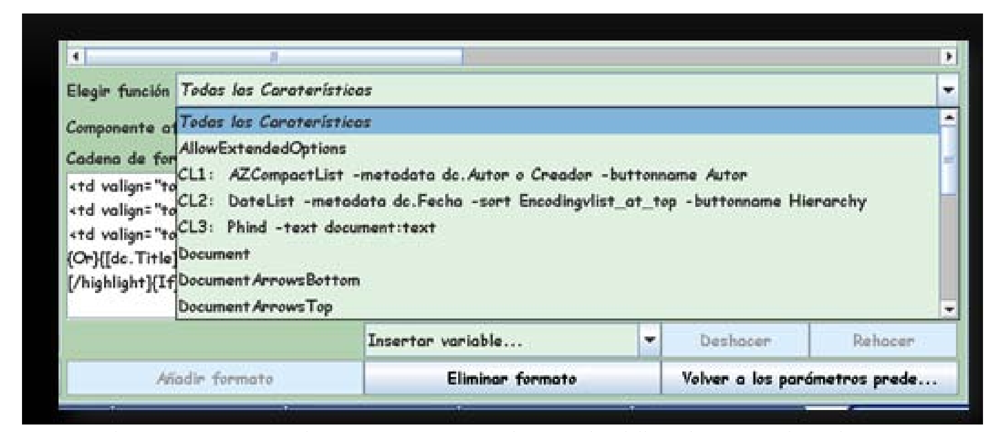


Figura 8.32 Opción Formato Incluye clasificadores y funciones.

El formato ***Vlist*** presenta los resultados en forma de lista vertical cuando se realiza una consulta.

La Figura 8.32 presenta las cadenas de formato dentro de la opción ***Formatos***, este controla la estructura y apariencia de los resultados de la búsqueda y la navegación y por consiguiente la forma en que aparecen los documentos en la colección.

Estos caracteres son metadatos que definen la configuración de sus colecciones llamadas ***cadenas de formatos*** (8). Según configure y dé las opciones en la ***cadena de formato*** así se visualizarán después los documentos en su página ***Acerca de...***

Puede aplicar una cadena de formato a cualquier componente que se encuentre en la lista llamada ***Elegir función***, (Ver Figura 8.32) la cual incluye cada clasificador y una lista de funciones determinadas previamente. Para cambiar una cadena de formato, selecciónelo de la lista, modifique sus parámetros y haga clic en ***Añadir formato***. Si desea insertar una nueva variable dentro de la cadena de formato, selecciónela de la lista descendente y haga clic en *Insertar*.

Cuando comienza a trabajar con estas cadenas, resulta un poco complicado seleccionarlas adecuadamente; pero en la medida en que practique el lenguaje de formateo y logre el efecto de la apariencia de las colecciones deseado por usted, se convierten en una poderosa herramienta de trabajo. Para agilizar este proceso puede crear un documento *.doc o .txt* en la que guarde las formadas por usted y con estas mismas, (si por ejemplo cuando revisó su colección le satisfizo la forma en que se presentaron los documentos), puede entonces establecerlas como “modelo” para trabajar posteriormente en la configuración de sus colecciones. Así para cada colección, si lo desea, usted copia y pega en el espacio de cadenas de formato estas establecidas como “modelo”. De esta forma quedarían según su gusto y forma de organizar la información teniendo en cuenta, siempre, las necesidades de los usuarios finales.

Greenstone dispone de valores predeterminados para cada cadena de formato, diseñadas para producir algo sensato bajo cualquier condición; pero usted puede configurar cadenas nuevas que modifiquen el resultado de la apariencia de los documentos dentro de sus colecciones e invalidar las que genera por defecto el programa. Por eso es necesario hacer muchas pruebas hasta lograr la más adecuada a su gusto.

El uso estándar de instrucciones de formato se aplica a resultado de búsquedas, clasificadores, etc. Cuando se realiza una búsqueda la lista que se obtiene mediante la función ***Search*** se llama *Búsqueda*, y la lista generada por los clasificadores que usted seleccionó durante la configuración de la colección se denominan ***CL1***, ***CL2***, ***Cl3,*** etc, respectivamente (Ver Figura 8.32).

- VList se aplica a todas las listas verticales.
- HList se aplica a todas las listas horizontales.
- SearchVList lista vertical de los resultados de búsqueda.

Con el mecanismo ***[link]... [/Link]*** en una cadena de formato, se crea un enlace con el formato HTML del documento ***[link][Title][/link]*** donde el hipervínculo es el título del documento. Esta aparece en la primera línea de las cadenas si la elimina pierde el ícono de acceso al documento.

<td>[link][icon][/link]</td> Con esta cadena el ícono es el hipervínculo vinculado al texto. Estas dos son válidas para colecciones de documentos en formato HTML.

Aparecen etiquetas HTML entre signos **< >**, ejemplo **<td>,
, <i>...** y las propias de Greenstone entre corchetes y llaves Ej. ***([ex.sclink],{Or}, [ex.thumbicon]....***

Al combinar las etiquetas html, con las otras, obtendremos el formato que deseemos.

Para el **clasificador Azlist** Vlist aparece la siguiente declaración como formato predeterminado que se utiliza para cualquier lista vertical.

```
<td valign="top">[link][icon][/link]</td>
<td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
```

```
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
[/highlight]{If}{[ex.Source],<br><i>([ex.Source])</i>}</td>
```

Para el **clasificador AZCompacList** metadato autor, la cadena de formato predeterminada es:

```
<td valign="top">[link][icon][/link]</td>
<td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
[/highlight]{If}{[ex.Source],<br><i>([ex.Source])</i>}</td>
```

Para añadir un nuevo formato selecciónelo de la lista, modifique sus parámetros, cambie la información y haga clic en ***Añadir formato***. La nueva cadena de formato aparecerá en la lista de "Comandos o cadenas de formato asignados actualmente". Solamente se podrá asignar una cadena de formato a cada combinación de función y/o componente.

8.9.4- Ejemplos de etiquetas presentes en pestaña *Formato* opción Formatos

<hr> Etiqueta HTML, marca la línea de división.
<td> Etiqueta HTML para crear columnas y alinear elementos.
[ex.srclink] Etiqueta de Greenstone, crea un enlace al documento original de la colección.
<i> Etiqueta HTML, indica cursiva.
[dc.Title] Elemento de Metadato de Greenstone Dublin Core que corresponde al Título.
<strong> Etiqueta HTML, indica negrita
[dc.Creator] Elemento de Metadato de Greenstone Dublin Core que corresponde al Creador.
<p> Etiqueta HTML indica salto de línea
[dc.Description] Elemento de Metadato de Greenstone Dublin Core que se refiere a la Descripción.
<td valign=top>
<td > Es el elemento de inicio de una columna.
<valign>: Permite una alineación del texto en el sentido vertical de la celda.
<Top>: Alínea parte superior.

A continuación se representan las cadenas de formato asumidas por esta colección para cada función.

8.9.5- Ejemplos de cadenas de formatos para cada función y metadato

Función: CL1 AZ List metadata dc. Título; componente afectado: Vlist

```
<td valign="top">[link][icon][/link]</td>
<tdvalign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
[/highlight]
{If}{[dc.Creator],<br><i>Autor: [dc.Creator]</i>}
```

El resultado de la cadena de formato vista anteriormente se podría visualizar en la interfaz de título como aparece en la siguiente Figura 8.33.

CL1 AZList metadata dc. Título; componente afectado: Vlist En este caso el ícono que genera por defecto Greenstone es el que aparece en la parte superior izquierda de color azul, lo que significa que son documentos HTML y es esta instrucción ***[link][icon][/link]*** la que define que el ícono es el hipervínculo al documento.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

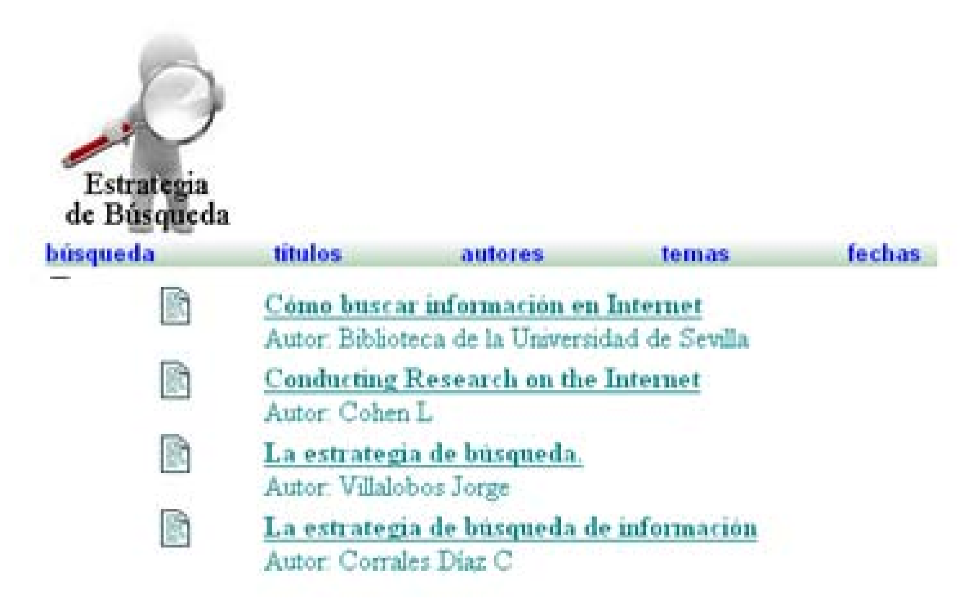


Figura 8.33. Ejemplo que representa la visualización de los documentos con CL1 AZList metadata dc. Título.

Otro ejemplo de cadena de formato sería para el clasificador *AZCompactList para metadato Autor*.

```
<td valign="top">[link][icon][/link]</td>
<td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
{If}{[dc.Creator],<br><b> Autor: [dc.Creator]</b></i>}
<td> {If}{[numleafdocs],<i>([numleafdocs])</i></small>}</td>
```

El resultado de la cadena de formato anterior se aprecia en la Figura 8.34



Figura 8.34 CL2 AZCompactList metadata dc. Autor o Creador; componente afectado: Vlist.

Si al final de la cadena agregamos la instrucción **[numleafdocs],** tiene el efecto de especificar entre paréntesis el número de elementos que se encuentran dentro de un estante, en el ejemplo anterior la cantidad de títulos que almacena cada autor.

Este clasificador genera estanterías que almacenan la cantidad de documentos por autor.

Para contar la cantidad de documentos agrupados

```
<td> {If}{[numleafdocs],<i>([numleafdocs])</i></small>}</td>
```

Numleafdocs: No será establecido para los documentos.

La instrucción **If** verifica el valor de la variable **numleafdocs**. Esta variable sólo se activa para nodos internos de la jerarquía y estanterías y suministra el número de documentos debajo de ese nodo.

Puede usted acceder a cada ícono de estantería con un solo clic para conocer los documentos que contiene cada autor tal como se distingue en la Figura 8.siguiente 35.



Figura 8.35. Visualización de los documentos almacenados en cada estantería

Para el metadato **palabras clave** con el botón de navegación al que hemos llamado **temas** se ha seleccionado el clasificador ***AZCompactList*** .

Función: CL1 3 ***AZCompactList*** y con la siguiente cadena de formatos se visualizan de la siguiente manera. Figura 8.36

```
<td valign="top">[link][icon][/link]</td>
<td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
```

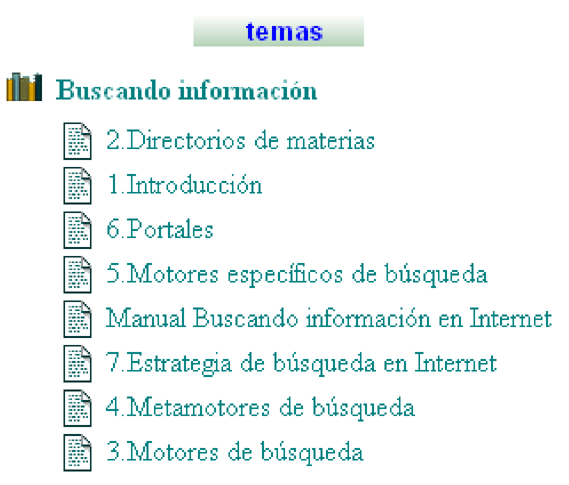


Figura 8.36 Visualización de documentos con clasificador AZCompactLIst

Para una cadena de formato para el C*lasificador Cl3 DateList metadata Fecha.*

```
<td valign="top">[link][icon][/link]</td>
<td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
[/highlight]
```

Se visualizan los documentos de la siguiente forma. Ver Figura 8.37

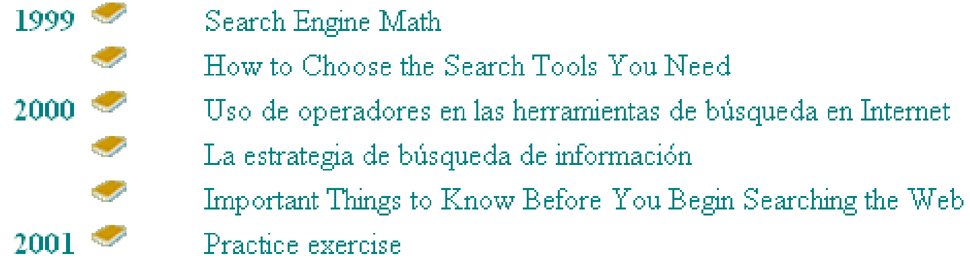


Fig 37 Listado de fecha según cadena de formato Cl3 DateList metadato Fecha

Donde el ícono de libro amarillo ( ) es el vínculo a los documentos.

Si desea crear una cadena de formato para procesar la función ***searchtype***, Ver Figura 8.38, esta permite la búsqueda por campos con una interfaz basada en formas, lo que indica además que se usa un motor de búsqueda avanzado como es MGPP.



Figura 8.38. Función searchtype

Figura 8.39.Resultado de la configuración de la función searchtype basada en formas.

En el ejemplo anterior se realizó la búsqueda a nivel del documento con la palabra *internet* en el campo título, aunque también puede hacerlo por los otros campos de búsqueda, el resultado final mostró 2 documentos dentro de la colección.

Estos ejemplos de cadenas de formatos demuestran que con pequeñas modificaciones y si utiliza diferente variables pueden lograrse cambios en los resultados finales para visualizar los documentos que conforman las diferentes colecciones digitales.

Es necesario enfatizar en el cuidado y esmero al crear o trabajar las cadenas de formatos; una simple omisión de alguna etiqueta, una llave, instrucción o signo que no se utilice correctamente afectaría ,indiscutiblemente, en el diseño de la apariencia de la página o no se lograría el efecto que se concibió originalmente en su diseño

8.10- Pestaña Crear

Después de haber definido una nueva colección, reunido sus documentos, y diseñado su apariencia, puede crear la colección (Figura 8.40).

Es recomendable enfatizar que no debe confundirse **Greenstone** con **GLI**. Esta última es solamente una herramienta que se puede usar para construir una colección con el sistema Greenstone.

En la opción ***Crear colección*** (Figura 8.40) con un clic comienza el proceso de creación de la colección, debe tener presente que cada vez que usted realice un cambio de índice debe crear de nuevo la colección, para que cargue o salve esos cambios, excepto en las macros específicas así como también en las cadenas de formato que no necesitan crearlas de nuevo; solamente con actualizar la página se visualizan inmediatamente los cambios efectuados.

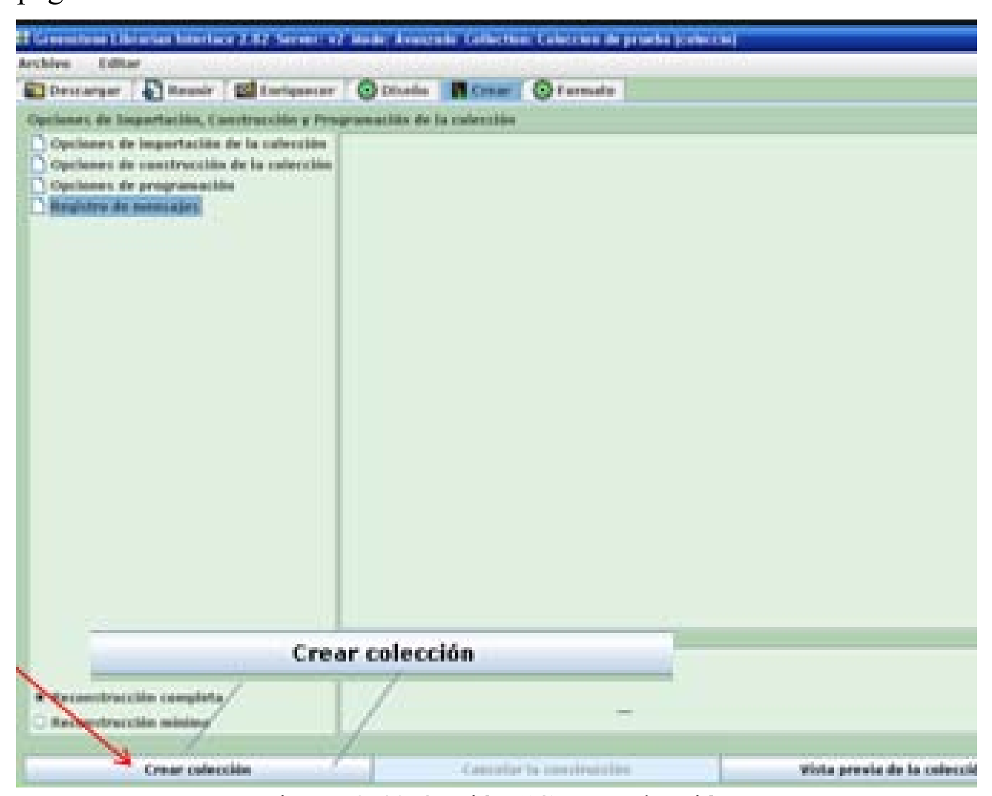


Figura 8.40 Opción “Crear colección

En la Figura 8.41 que se presenta a continuación muestra el proceso de importación y construcción de la colección; en la parte superior de la ventana, dos indicadores informan sobre el proceso en porcientos (%).

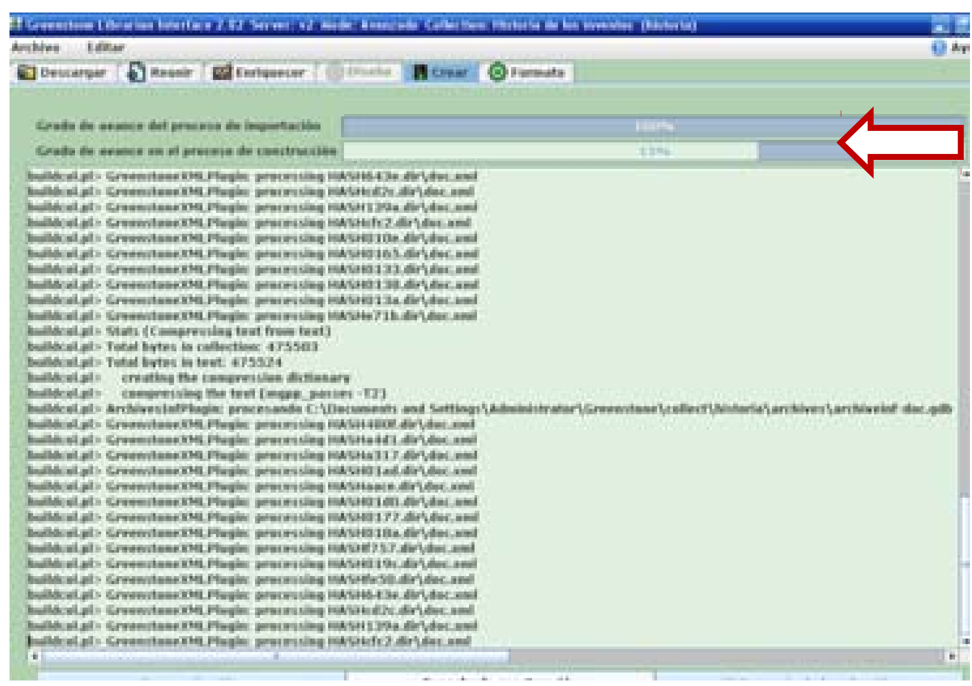


Figura 8.41 Proceso de importación y construcción de la colección

El desarrollo del proceso de importación y tratamiento de documentos lo muestra en el fichero de sucesos (log) y a la vez se despliega una serie de datos que le informan si los plugins que usted seleccionó en la pestaña ***Diseño*** reconocen los diferentes formatos de sus documentos, si fueron procesados correctamente y si serán vistos en la colección, la revisión muestra igualmente si algún documento ha sido rechazado, y le indica si los documentos se han importados correctamente y se convierten en lenguaje XML.

El usuario observa el proceso de construcción a través de una ventana que muestra no sólo los textos generados por los guiones de importación y creación de índices de Greenstone, sino también las barras de progreso que indican el grado general de avance de cada guión.

Resulta atinado aclarar que cuando usted observe algún documento rechazado tendrá la opción de cancelar la construcción de esta colección, observar las razones por las cuales el documento no ha sido procesado y ver si puede solucionar la situación, realice los cambios pertinentes y recomendados, guardar el cambio y de nuevo dar un clic en ***Crear.***

Consecutivamente aparece una ventana automática como la que se muestra en la Fig.42 avisándole que la colección ha sido creada. Este mensaje le indica que la colección está construida exitosamente y que está lista para verla.



Figura 8.42 .Resultado del proceso de creación de la colección

Puede dar un clic en la opción ***Vista previa,*** de este modo enseguida carga la colección para que tenga una visión de cómo ha quedado su colección creada.

Cuando el usuario da un clic en un ícono, aparece la página ***Acerca de*** de la colección correspondiente. Cada vez que se carga la página, se genera dinámicamente el menú de las colecciones basándose en las colecciones que se encuentran en el sistema de archivos en ese momento. En este caso ya con una colección creada y con el ícono que la representa.

Ahora utilice el navegador para acceder a su biblioteca, y consulte la colección formada con el título: "Estrategia de búsqueda".



Figura 8.43 Visualización de la página de la colección creada

Colección Estrategia de búsqueda

Descripción de esta colección

Buscar información en Internet representa una de las habilidades indispensables a obtener en el desarrollo exitoso de una Sociedad de la Información -Conocimiento. La labor desplegada hasta muy recientemente por el personal especializado de la biblioteca pasa a ser en nuestros días una metodología de necesaria incorporación por todos aquellos que requieran localizar información pertinente y oportuna.

En esta colección Ud. podrá disponer de una información detallada sobre todos los aspectos y características de esta estrategia de tan valiosa incorporación en la sociedad digital en la que ya nos encontramos inmersos.

Una vez finalizado el proceso de creación de la colección y comprobado los resultados deseados, puede usted cerrar la interfaz del bibliotecario por la opción ***Salir*** del menú ***Archivo*** tal como se observa en la Figura 8.44



Figura 8.44 Opción Salir para cerrar GLI.

Proponemos al final de este capítulo un instructivo que sintetiza de manera práctica los aspectos clave para crear una colección digital con el programa Greenstone de forma eficiente.

8.11- Consideraciones finales

El ejemplo desarrollado se fundamenta en la construcción de una colección con documentos HTML con características precisadas, resultado de una práctica por medio de la interfaz GLI con todas las facilidades de acceso que entrega el sistema Greenstone. Este prototipo de colección ofrece un poderoso mecanismo de organización que garantiza la óptima recuperación de la información con una apariencia interesante y atractiva para convertirlo en recurso privilegiado en beneficio del usuario, satisfacer sus necesidades y logre resolver perfectamente los problemas básicos de recuperación de información textual finalmente.

La flexibilidad ofrecida por Greenstone resulta fundamental para lograr una interfaz con una apariencia en la que se exprese la creatividad del bibliotecario que demanda un usuario final y pueda aprovechar todo el potencial que brinde la colección consultada.

La habilidad para crear los formatos de salida y de visualización de documentos requiere de

opciones avanzadas y de conocimientos por parte del creador, es necesario conocer con detalles parámetros y metadatos disponibles para cada colección.

8.12- Referencias bibliográficas

(1) Mendez E. Senso A. Introduccion a los metadatos. Estándares y aplicación. Uso del Dublin Core. Unidad de Autoformación.[Sitio Web en Internet] SEDIC 2004. Disponible en: <http:www.sedic.es >

(2) Vera F. Barrera R. Herramienta gráfica GLI; gestión y desarrollo. En Curso avanzado para la construcción de bibliotecas digitales utilizando el software Greenstone. La Habana CECAM: Prodigio consultores ,2009.Dublin Core Metadate Iniciative.

(3) Conjunto de Elementos de Metadatos Dublin Core, Version 1.1: Descripción de Referencia. [Sitio Web en Internet]. 2003-06-02. Disponible en <http://es.dublincore.org/documents/2003/06/02/dces/index.shtml.rdf >

(4) Tramullas J. Garrido .Software libre para gestión de recursos de información digital. Disponible <http://tramullas.com >

(5) Tramullas J. Garrido P. Bibliotecas Digitales: Greenstone. Cap X. En: Software libre para servicios de información digital. Madrid: Pearson Prentice Hall .2006. [Libro en Internet]. Disponible en < http://eprints.rclis.org/6145/1/Capitulo10.pdf >

(6) Bainbridge D, McKay D, Witten H. I. Guía del programador. Biblioteca digital Greenstone. [Sitio Web en Internet].Departamento de Informática. Universidad de Waikato, Nueva Zelanda. Marzo de 2004.Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Develop_es_all.html >

(7) Masternagazine. Definición de Plug-in.[Sitio Web en Internet] 16 de Febrero de 2005. Disponible en: http://www.mastermagazine.info >

(8) Álvarez M.A. Documento introductorio al lenguaje Perl. .[Sitio Web en Internet] 29/9/01. Disponible en: <http://www.desarrolloweb.com/contact/1.htm > 29/9/01

(9) Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 2. Explotación Óptima de sus Documentos. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/style.css >

(10) Vera F. Desarrollo de bibliotecas en Chile utilizando Greenstone y su vinculación en la cooperación regional. En: Greenstone: Un software libre de código abierto para la construcción de bibliotecas digitales.2010, p.75-86. UNESCO. Montevideo. [Libro en Internet]. Disponible en: <http://www.unesco.org.uy/ci/fileadmin/comunicacion-informacion/Greenstone-paraweb.pdf >

(11) Witten IH, Boddie S, Thompson J. Guía del usuario. Nueva Zelanda: Universidad de Waikato, 2006.Disponible en <http://www.greenstone.org/manuals/gsdl2/es/html/User_es_index.html >

CAPÍTULO 9

9- Preparación de PDF para la construcción de una colección Greenstone

Dr. C. Raul G. Torricella Morales (torri@reduniv.edu.cu). Prof. Titular adjunto, Facultad de Farmacia y Alimentos de la Universidad de La Habana, Director de la Editorial Universitaria (Cuba). Dirección: Calle 23 # 525 Esq. F., Municipio Plaza, Ciudad Habana, Cuba, CP: 1400, Telf.: (537) 830-3674. Ministerio de Educación Superior de la República de Cuba.

Palabras clave: Repositorio Institucional, Recursos educativos en línea, Biblioteca digital, Biblioteca Virtual, Documentos académicos, Editoriales universitarias.

9.1- Introducción

El formato PDF se utiliza ampliamente para la divulgación y publicación de documentos académicos y científicos, por lo cual este formato es uno de los más utilizados en la construcción de Bibliotecas digitales y repositorios institucionales y temáticos. Sin embargo, muchos de estas colecciones no realizan ninguna preparación de los mismos, lo que afecta su usabilidad.

En este capítulo, se discuten algunos aspectos teóricos y prácticos que deben tenerse en cuenta antes de comenzar a construir una colección de PDF, para lo cual emplea el caso de estudio de la Biblioteca Virtual de la EcuRed (Repositorio Institucional del Ministerio de Educación Superior) disponible en Internet en: http://revistas.mes.edu.cu

9.2- Socialización del conocimiento: acceso abierto y privativo

La Iniciativa de Acceso Abierto parte del principio de que se depositen en Internet los documentos científicos (o de cualquier otro tipo) para su: lectura, descarga e impresión de forma libre y gratuita (1, 2).

Entre las principales características que se deben tener en cuenta a la hora de implementar un proyecto de biblioteca en la modalidad de Acceso abierto, están las siguientes:

- Establecer el tipo de licencia para el acceso a la información de manera que se beneficien tanto los autor@s como los lector@s.
- Implementar las características técnicas de los formatos y estructuras de comunicación compatibles con los protocolos de almacenamiento e intercambio de información con y entre otros repositorios de Acceso Abierto y de Revistas Open Access.
- Tener en cuenta las cuestiones legales y de financiamiento de las publicaciones.

El número de usuarios en Internet en el año 2000 era de 357 millones, ya en el 2010 se acercaba a los dos mil millones[4]. Este salto cuantitativo, unido a el impetuoso desarrollo de las tecnologías de la información y las comunicaciones, ha logrado que el proyecto de Acceso Abierto a las

[4] Son 1 967 millones de usuarios según: Adama Web S.L. (2011). Difusión y divulgación científica en Internet. España: Gobierno del Principado de Asturias. – Pág. 12.

publicaciones sea ya una práctica común (3).

Un ejemplo de esto se tiene en España, donde es obligatorio publicar en la modalidad de Acceso Abierto (OA):

- Las Publicaciones financiadas con fondos públicos europeos.
- Las Tesis doctorales realizadas en Universidades españolas.

Por otra parte, la UNESCO promueve otra iniciativa semejante, pero para los materiales didácticos: Los Recursos Educativos Abiertos (REA) o por su nombre en inglés *Open Educational Resources* (OER) (4).

Los REA se definen (4) como medios de enseñanza-aprendizaje con licencia de "*dominio público*" o con alguna otra de Acceso Abierto que permita su uso, re-ensamblado, reutilización y redistribución, por terceros sin restricción alguna. El empleo de normas técnicas permite el intercambio de información con otros repositorios. Entre los recursos educativos se encuentran los libros de texto, las guías de estudio y las notas del profesor, estos últimos se clasifican como e-Libros de Texto (e-Textbook) o simplemente e-Textos.

Los e-Textos constituyen (5) los recursos educativos tradicionales de mayor importancia en el proceso de enseñanza-aprendizaje, tanto en la educación pre-universitaria como la universitaria, sin embargo, aún muchos autores no están dispuestos a permitir la distribución de sus obras con licencias del tipo Acceso Abierto.

Los e-Textos se distribuyen en redes privadas o intranet institucionales, para uso exclusivo de las universidad que los generó, o se distribuyen a través de editoriales comerciales que pagan "derechos de autor" o compran los derechos de publicación a los autores de las obras.

La distribución en la modalidad de "*Socialización del Conocimiento*" (5) permite que los contenidos puedan accederse tanto en Acceso Abierto como comercial, siempre y cuando esta última modalidad agregue valor y pague los derechos de copia a quien los ostente. La comercialización exige el empleo de tecnología apropiada para restringir las copias no autorizadas.

Este tipo de distribución se realiza mediante plataformas tecnológicas comerciales, como por ejemplo: e-libro. Esta plataforma brinda acceso a la producción de la Editorial Universitaria, que también tiene su producción en Acceso Abierto en la Biblioteca de la EcuRed (http://revistas.mes.edu.cu), le añade valor y le incrementa la visibilidad a estos contenidos, además de que le permite a los autores de esta editorial el acceso abierto a los recursos disponibles en la plataforma.

En este sentido, las soluciones tecnológicas que se emplean en la producción de contenidos para su comercialización en línea, pueden y se recomiendan utilizar en la producción de contenidos destinados a su divulgación en la modalidad de acceso abierto. En este capítulo se presentan técnicas para la preparación de documentos para ser incorporados en una biblioteca digital con el uso del software Greenstone, para lo cual se parte de los procedimientos que se emplean en la construcción de la Biblioteca Virtual de la EcuRed.

9.3- Biblioteca Virtual de la EcuRed

Algunos autores utilizan el término "*biblioteca virtual*" en lugar, o como sinónimo, de "*biblioteca digital*", (6) aunque es más correcto es este último, en este documento se mantiene la denominación original por la cual se conoce este recurso: *Biblioteca Virtual de la EcuRed*, y está disponible en Internet la dirección http://revistas.mes.edu.cu y es uno de los cuatro repositorios cubanos que están

registrados en el Directorio de Repositorios de Acceso Abierto[5] (DOAR).

Esta biblioteca divulga la producción de la Editorial Universitaria (cubana) así como una selección de los e-Libros publicados por algunas editoriales del Instituto Cubano del Libro que han facilitado sus contenidos su socialización y así servir de fuentes para la construcción de artículos enciclopédicos en la EcuRed (7).

El contenido de esta biblioteca se selecciona sobre la base de su calidad y utilidad como recurso educativo para las universidades cubanas. Por supuesto que los autores de los e-Libros de texto deben dar su conformidad para distribuir sus obras bajo la licencia de Acceso Abierto y en plataformas comerciales, como una vía adicional para el incremento de la visibilidad de los contenidos de la Biblioteca. La licencia que emplea la Editorial Universitaria es la siguiente:



Creative Commons de tipo Reconocimiento, Sin Obra Derivada, se permite su copia y distribución por cualquier medio siempre que mantenga el reconocimiento de sus autores y no se realice ninguna modificación de ellas.

La Figura 9.1 muestra el modelo conceptual subyacente de la gestión de los contenidos que realiza la Editorial Universitaria para la divulgación de los e-libro en la Biblioteca Virtual de la EcuRed mediante el uso del software Greenstone 2.84.



Figura 9.1 Modelo conceptual de la gestión de información que realiza la Editorial Universitaria cubana para incorporar documentos en la Biblioteca de la EcuRed y el Repositorio de recursos educativos del Ministerio de Educación Superior[6].

Los documentos que se incorporan, en función de los "Derechos Digitales" que se acuerden con los autores o tengan declarados los documentos, en la biblioteca de la EcuRed (acceso abierto) o en la intranet (acceso restringido).

Las fuentes de obtención de los recursos digitales pueden ser una de las siguientes:

- Documentos disponibles en la Web en formato de imagen.
- Documentos disponibles en la Web en formato PDF
- Producción documentos académicos a partir de originales.

A continuación se describen los procedimientos para "*Preparación de los documentos para la construcción de bibliotecas digitales*" soportadas en Greenstone"

9.4- Documentos disponibles en la Web en formato de imagen

En la Web se pueden obtener muchos libros de texto en formato imagen o en PDF generados a partir de imágenes, los cuales "no tienen fuentes", el texto es una imagen y como tal son de muy poca utilidad para construir biblioteca digitales: por una parte son ficheros de gran tamaño y por la

[5] Véase DOAR en: http://www.opendoar.org/

[6] Figura elaborada a partir de la lámina 29 de la presentación: Vera L., Felipe (2006) Construcción de bibliotecas digitales: modelo de capa piramidal. Ponencia presentada enL XI Conferencia Internacional de Bibliotecología: "Bibliotecología: Oportunidades y desafíos en la Sociedad del Conocimiento". Santiago de Chile, 2006: pp. 29

otra no se puede buscar el texto dentro del documento.

Para convertir estas imágenes a texto debe emplearse algún "*sistema de reconocimiento óptico de caracteres*". Uno de los más utilizados es el ABBYY FineReader 8.0 Profesional Edition (8). A continuación se describen las principales operaciones para convertir las imágenes en ficheros PDF nativos, a manera de caso de estudio. Si Ud. requiere la realización de este tipo de conversión, le recomendamos que adquiera el software o busque una alternativa en software libre[7].

El ABBYY FineReader puede crear tres tipos de ficheros PDF:

- Sólo texto e imágenes.
- Texto bajo imágenes.
- Texto sobre imágenes.

La primera opción produce los PDF de mayor calidad, pero es la más laboriosa, pues todos los errores de conversión del texto son visibles para el lector, lo que hace necesario su total corrección. Esto suele ser muy laborioso, sobre todo si las imágenes originales no son de calidad. Las imágenes decorativas, como son líneas a la cabeza o al pie de las páginas, deben marcarse todas para que se incorporen en el documento.

Cuando se tienen diseños complicados, viñetas especiales, líneas y otros adornos que no se quieren perder y el fondo está relativamente limpio, se recomienda la utilización de la opción "Texto sobre imágenes". En este caso no es necesario marcar las líneas, estas aparecerán en los texto, pero también las imperfecciones de la digitalización, las cuales deberán "borrarse" de los originales, para lo cual se cuenta con una herramienta específica para ello.

Cuando se quiere mantener el efecto original del texto, o la calidad de las imágenes es muy baja, se recomienda el empleo de la opción "Texto bajo imágenes": el lector ve la capa de imagen con su calidad original, y por debajo hay una capa de texto que puede buscarse, copiar y pegar, aunque los errores no se ven "a primera vista" siguen presentes y afectan el valor de uso del documento.

Figura 9.2 Resultados del reconocimiento de texto con la opción "Sólo texto e imágenes" en documentos con baja calidad. A la izquierda la imagen original, a la derecha los resultados del reconocimiento automático.

La Figura 9.2 presenta el aspecto de un documento relativamente antiguo (el documento original de la Reforma Universitaria en Cuba, con 50 años de existencia y digitalizado con muy baja calidad)

[7] Aclaración del autor: Estamos en contra de la Piratería, pero tampoco estamos promoviendo el uso de software propietario. Sólo se emplean las referencias al ABBYY Finereader como un ejemplo metodológico.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

en su versión original y después de ser reconocido con la opción "Sólo texto e imágenes", sin correcciones.

Como se puede apreciar en la Figura 9.2 las imágenes de baja calidad tienen un efecto negativo sobre la eficiencia del reconocimiento de los textos, a tal punto que en algunos casos sería menos costoso reescribir los textos que someterlos al proceso de reconocimiento automatizado mediante algún software.

A continuación se presenta como ejemplo, la referencia del libro, que actualmente está disponible en la Biblioteca Virtual de la EcuRed:

Evaluación Sensorial aplicada a la investigación, desarrollo y control de la calidad en la Industria Alimentaria / Raúl G. Torricella Morales, Esperanza Zamora Utset y Horacio Pulido Alvarez. La Habana : Editorial Universitaria. -- 2008. -- 136 pág. 1,9 Kb.

Se reescribió completamente a partir de su original impreso, que aunque data sólo del 1989, su calidad era insuficiente, y clasifica como un "Texto Robotron", o sea que el texto está escrito con una máquina de escribir mecánica, con caracteres no proporcionales, lo dificulta aún más el reconocimiento automático.

Siempre que sea posible, es mucho más conveniente utilizar documentos disponibles en formato PDF nativo, es decir, que han sido creados a partir de algún procesador de texto, como el Word, LibreOffice u otro similar. En estos casos es necesario realizar un conjunto de operaciones para que estos PDF puedan ser procesados convenientemente por el Greenstone para su incorporación en una colección. Estas operaciones se describen a continuación.

9.5- Documentos disponibles en la Web en formato PDF

La Editorial Universitaria cubana emplea el procedimiento que exige la plataforma e-libro para la incorporación de contenidos en ella. De esta formato se cumplen dos objetivos con el esfuerzo de uno: se preparan los documentos para la Biblioteca Virtual de la EcuRed y para su entrega a la plataforma e-libro, cuando esto proceda. A continuación se presenta una versión simplificada de este procedimiento.

9.5.1- Producción de documentos PDF para bibliotecas digitales

Una vez se identifica un documento en PDF, que por supuesto no está aún en la biblioteca, se procede a revisar su calidad. En primer lugar, el documento debe ser pertinente para la colección a la cual se destina. Los documentos que cumplen con estos requisitos y son aceptados, se incorporan al flujo de trabajo.

La primera acción es verificar si el PDF tiene algún tipo de seguridad activada; si la tiene, se debe quitar, para lo cual se puede utilizar algún programa diseñado al efecto, por ejemplo el *Advanced PDF Password Recovery Pro*. (Proceso 1 en la Figura 9.3)

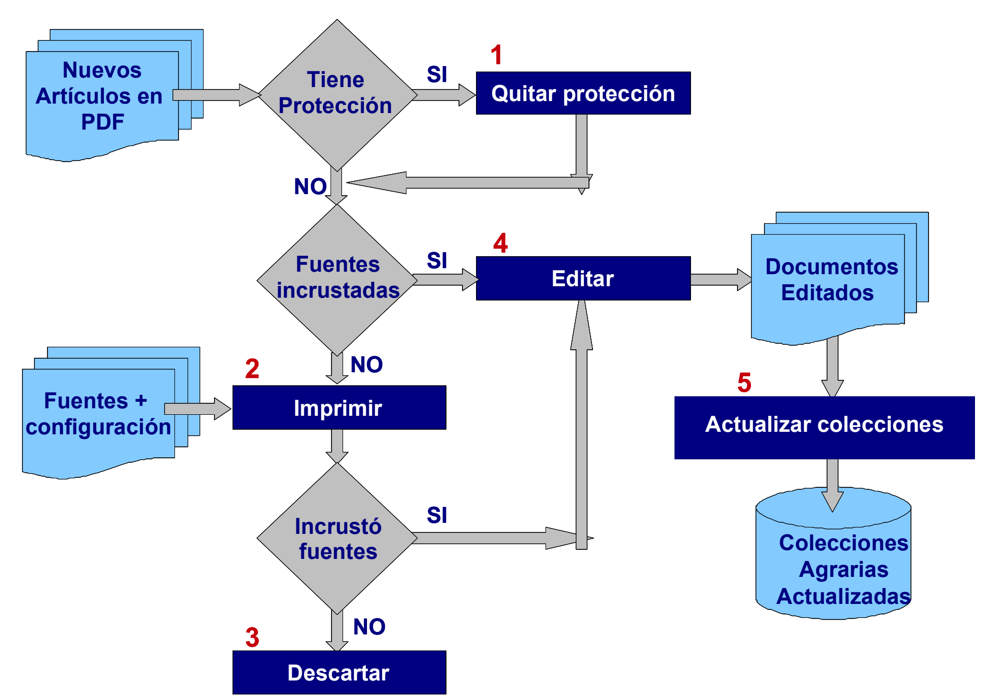


Figura 9.3 Flujo de trabajo para producir documentos en PDF para una biblioteca digital en Greenstone (Tomado de: Cruz Santos, Ileana Dayamina de la, 2011, pág. 73.)

Luego se comprueba que los ficheros PDF tengan todas las fuentes incrustadas. Se admite que las fuentes básicas del Adobe no estén incrustadas. (véase Tabla 9.1).

Tabla 9.1 Fuentes básicas del Adobe Acrobat

Courier	Helvetica-Oblique
Courier-Bold	Symbol
Courier-BoldOblique	Times-Bold
Courier-Oblique	Times-BoldItalic
Helvetica	Times-Italic
Helvetica-Bold	Times-Roman
Helvetica-BoldOblique	ZapfDingbats

En el caso de que tenga alguna fuente no incrustada, se procede a imprimir el documento con la configuración de impresión: *eBook.joboptions*[8]. (Proceso 2 en la Figura 9.3). Hay que tener en cuenta que los artículos deben tener todas sus páginas de un mismo tamaño (desde la tapa -o cubierta- hasta la última página). Para comprobar la eficacia de la impresión se elige una porción de texto al azar, seleccionar, copiar y pegar en un fichero de texto. Si la copia se puede leer correctamente, se continua la edición, si lo que se copió son símbolos ilegibles, o está en blanco, entonces el artículo se descarta (proceso 3 en la Figura 9.3).

La edición (proceso 4 en la Figura 9.3) se realiza a partir del PDF impreso, o de una copia (cuando todas las fuentes están incrustadas). El original se guarda en una carpeta de “originales” para asegurar que se mantenga inalterable. Se debe comprobar que el artículo esté completo y hacer que coincidan los números de páginas en las miniaturas con los números reales de las páginas. A la primer página siempre se le llamará *Cover*. En el caso de que haya páginas en blanco se comprueba la razón por la cual están en blanco; si es para hacer coincidir la numeración, entonces se las deja, de lo contrario se eliminan.

[8] Esta opción está definida en un fichero de configuración disponible en la dirección http://bives.mes.edu.cu

El proceso de edición continúa con la escritura en los campos “Título” y “Autores”del PDF de los siguientes datos:

En el **título**: Título completo del artículo, si es mayor de 250 caracteres se trunca. A continuación punto, el nombre o la abreviatura de la publicación seriada, el número del volumen, el número entre paréntesis, dos puntos, página inicial y final separados por un guión, coma y el año de publicación. Por ejemplo:

- Detección de antibióticos en leche comercializada en la Ciudad de México. Rev. Salud Anim. 23 (1):37-41, 2001

En el campo **Autor** completar con todos los autores que figuren con el formato: Apellido/s, Nombre/s; Apellido/s, Nombre/s... etc. Si tuvieran otro rol que no fuera de Autor, como por ej coordinador, editor, etc. Entonces se aclararía esto entre paréntesis después del nombre de c/u. Por ej: Apellido, Nombre (Autor); Apellido, Nombre (Coordinador). Si son todos autores no se aclara, es decir no se pone (Autor), solo se aclara en el caso que se esté aclarando otro de ellos. Por ejemplo:

- "Ramírez, Acacia; Gutiérrez, R.; González, Clementina"

A continuación se modifican las opciones de apertura del PDF de la siguiente forma:

- página y panel de marcadores
- una sola página
- ajustar ancho
- y que inicie desde la primer página (Cover)

Finalmente se procede a crear los *bookmark* (marcadores en español) de la siguiente forma:

- El primer bookmark siempre es el título, y lleva al principio de la primer página o cover del artículo.
- El segundo es uno que se llamará “Página Legal” y llevará a una página en donde figuran todos los datos legales del libro, como ISBN o ISSN, Editorial, etc (si es que tiene esta página, ya que en el caso de los artículos no está, por lo que no se hace).
- El tercero es el índice (si existe). Los artículos generalmente no tienen índice, por lo que no se agregará.
- Y después se hará mínimo un bookmark por cada título/subtítulo/capítulo, etc. que tiene el artículo. Debe posicionarse en cada página donde se encuentra un título, subtítulo para crear el bookmark.
- Una vez creados todos, se acomodan como un árbol, cuya rama principal es el nombre del documento y de ahí hacia abajo los otros títulos, capítulos, y subtítulos.

A continuación se presenta ejemplo de la estructura de los bookmarks y el uso correcto de las mayúsculas y minúsculas según el nivel de profundidad.

Tabla 9.2 Ejemplo de construcción de bookmarks

Correcto	Incorrecto
TÍTULO - SUBTÍTULO1 - SUBTÍTULO2 - SUB-SUBTÍTULO 1	Título 1 - Subtítulo1 - SUB-SUBTÍTIULO2 - Subtítulo2
TÍTULO - SUBTÍTULO1 - SUBTÍTULO2 - Sub-Subtítulo 1	TÍTULO 1 - Subtítulo1 - Sub-subtítiulo2 - SUBTÍTULO2

Los vínculos siempre deben tener las propiedades de Tipo de vínculo: invisible y Estilo de resaltado: ninguno. También deben hacerse los vínculos Web a las direcciones de e-mail (se hacen: mailto:dirección de e-mail, por ejemplo: mailto:info@e-libro.com) y las direcciones Web (http://dirección).

Al finalizar con la edición, grabar siempre el documento con la opción "guardar como..." de esta manera se eliminan todos los contenidos duplicados, por lo que se reduce el tamaño del PDF.

Si está trabajando con una versión de Adobe Acrobat superior a la 7.0, cuando se "guarda como" se selecciona la opción *Adobe PDF Files, Optimized (*.pdf)* y en la parte derecha se especifica la versión con la cual el PDF será compatible para que el Software Greenstone (versión 2.84) lo pueda abrir, se da click en el botón pincha en OK y luego en el botón guardar.

Una vez terminado, revisar el PDF, ver con las propiedades que se abre, probar unos bookmarks, verificar al azar algunos *links* (enlaces) para ver que todo haya quedado correctamente, se debe revisar todo el PDF, de arriba a abajo, para verificar que todo esté correcto.

9.6- Producción de documentos académicos a partir de originales

9.6.1- Tesis de grado, maestría y doctorado

En el Ministerio de Educación Superior, a partir del curso 2008/2009, se establece como un requerimiento para la realización de la defensa de los trabajos de grado (en opción al título de Licenciado o Ingeniero), la entrega del documento de tesis en formato digital. Con el objetivo de facilitar a los estudiantes y bibliotecarios la producción, almacenamiento y recuperación de las tesinas, Torricella Morales y cols. (9) desarrollaron un procedimiento formal y una propuesta concreta de procedimiento, la cual se implementó, en el Departamento de Alimentos del Instituto de Farmacia y Alimentos de la Universidad de La Habana (IFAL) durante tres años consecutivos con resultados satisfactorios.

Una parte importante del procedimiento la constituye la producción de las tesinas a partir de plantillas desarrolladas al efecto. Las plantillas facilitan a los estudiantes la construcción de la estructura capitular, los formatos de tablas y figuras y la inclusión de los datos bibliográficos y de licenciamiento, de manera que el fichero en formato PDF que se obtiene como producto se puede utilizar directamente para ser incorporado en la biblioteca digital.

El procedimiento propuesto establece que el estudiante es quien escribe y edita la tesis, lo que lo convierte en *Autor-Editor* del documento, mientras el tutor,es *coautor-editor*. De los siete actores que participan en todo el proceso que se propone (véase Tabla 9.1), el. *responsable del trabajo de investigación estudiantil* es el más importante. Es él quien, en su doble función de controlar la calidad de los documentos y el funcionamiento del proceso, asegura que las tesis se recopilen e incorporen en la biblioteca. Torricella y cols. (9) demostró que allí, donde no se cuenta con este

actor, no se logra producir, y mucho menos mantener una biblioteca digital de tesinas.

Por otra parte el bibliotecario es quien construye la biblioteca a partir de los documentos digitales elaborados por los estudiantes y tutores; por eso, todos deben trabajar en estrecha colaboración bajo la dirección de los responsables de investigación estudiantil, quienes conjuntamente con los bibliotecarios deben impartir seminarios de capacitación para el uso de las plantillas y la producción de documentos que respondan a las exigencias que se requieren para su incorporación en la biblioteca digital.

En la Tabla 9.3 Se presentan las relaciones entre los principales actores del procedimiento para la producción y divulgación de tesinas en la biblioteca digital.

Tabla 9.3. Descripción de las principales relaciones entre los los actores del procedimiento que se propone para la divulgación de las tesinas en la biblioteca digital (9)

Actor	Descripción de las relaciones de los actores expresadas en sus acciones realizadas durante el proceso de escritura, revisión y edición de las tesinas
Autor-editor	**Escribe** la tesina con el empleo de **plantillas** de **LibreOffice** 3.3 con la ayuda del gestor de referencias bibliográficas **BiblioExpress** Ver. 2.0 y lo salva en formato PDF.
	Entrega copia de la tesina (en PDF para su revisión y en ODT para su digitalización) y de su biblioteca digital en soporte tangible (por ejemplo en un CD-ROM) al responsable del trabajo de investigación estudiantil.
Tutor	**Escribe** su opinión .y verifica que el autor la incluya en el cuerpo de la tesina.
	Propone la modalidad y el momento de la publicación de la tesina.
Responsable del trabajo de investigación estudiantil	**Revisa** el contenido del CD y el formato: si cumple lo establecido, lo recibe, en caso contrario, lo devuelve al autor para que lo rectifique.
	Copia los ficheros en ODT, PDF y las bibliotecas personales de los autores en en carpetas debidamente identificadas y codificadas para uso interno del Departamento docente. Esta copia forma parte del depósito legal institucional, conjuntamente con el CD que será entregado a la biblioteca de la institución.
	Entrega al oponente el CD y le hace firmar el acta de recepción con el compromiso de devolverlos una vez concluida la defensa del trabajo de diploma.
	Recolecta y revisa los CD y se asegura que se corresponda con la copia en PDF que obra en su poder.
	Entrega al bibliotecario copia en soporte CD de las tesinas defendidas y aprobadas por el tribunal para para ser archivados en el depósito legal institucional.
Oponente	**Recibe** el CD con la tesina para su revisión, escribe su opinión y participa, conjuntamente con el tutor y el tribunal, en la definición de la modalidad y el momento de publicación de la tesina. El CD se entrega a la biblioteca de la institución para formar parte del depósito legal institucional.

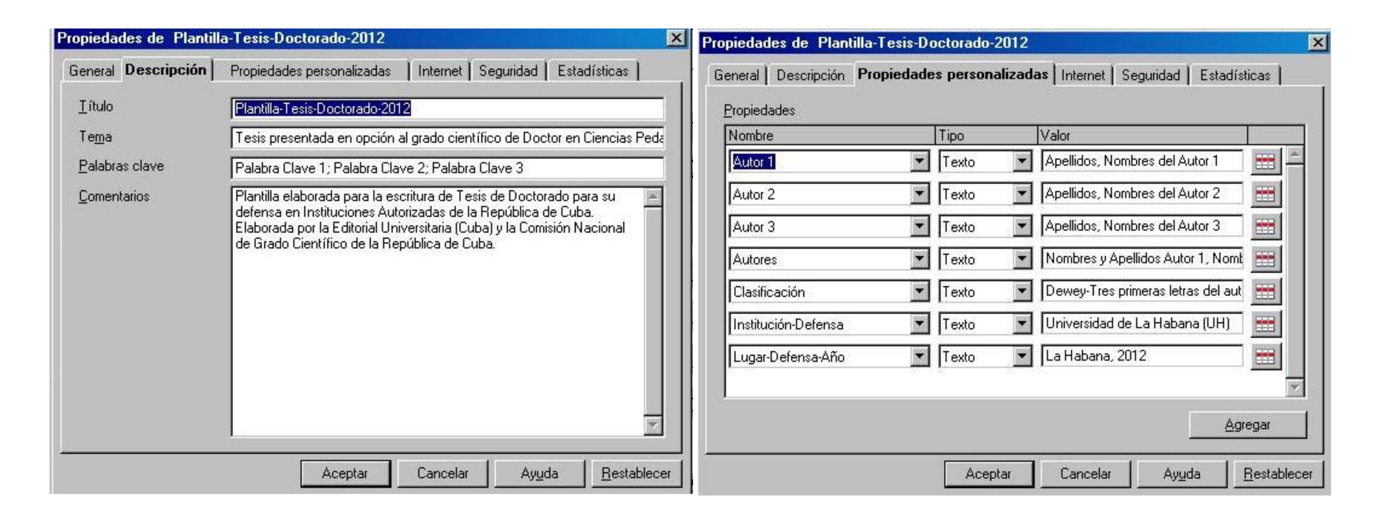
Actor	Descripción de las relaciones de los actores expresadas en sus acciones realizadas durante el proceso de escritura, revisión y edición de las tesinas
Bibliotecario	**Recibe** copia, en formato PDF, de las tesinas y las instrucciones para su divulgación.
	Registra los metadatos de las tesinas en la base de datos bibliográfica de la biblioteca y en el repositorio de la universidad.
	Sube al repositorio los ficheros en PDF de las tesinas, aprobadas para su divulgación.
Director de la Red	**Asesora** a estudiantes, profesores y bibliotecarios en el proceso de digitalización.
	Asegura que el repositorio de tesinas esté visible en la intranet e Internet según corresponda.
Director del repositorio	**Mantiene** el correcto funcionamiento del repositorio de tesinas.
	Forma a los bibliotecarios para procesar las tesinas.

La plantilla se elaboró en *LibreOffice* ver. 3.3 para *Windows,* pero también funciona con LibreOffice sobre linux. Constituye un modelo para quienes quieran construir sus propias plantillas a partir de la adaptación de una ya elaborada. Pueden descargar la plantilla en la dirección: http://bives/mes.edu.cu/install/Plantilla-Tesina-IFAL.odt A continuación se describe las principales características que debe cumplir las plantillas para la escritura de los trabajos de diploma.

9.6.1.1- Metadatos o descripción del documento de tesis

La plantilla tiene varios componentes, los relacionados con los "*metadatos o descripción"* del documento. Estas son las que se definen e incorporan en el documento por el usuario en los campos correspondiente, y luego pueden insertarse en las páginas que correspondan, por ejemplo: portada, portadilla, página legal (o créditos), pie de página y encabezamiento de página. De esta forma se asegura que siempre aparezcan escritas correctamente y que sólo sea necesario escribirlas una vez.

Para establecer las propiedades del documento abra el menú desplegable "*Archivo*" (primero en la parte superior izquierda), seleccione "Propiedades" (se abre una ventana) y dentro de esta las pestañas "Descripción del documento"y luego "Propiedades personalizadas". En primera pestaña hay tres campos: título, tema y palabras claves, en la segunda encontrará una tabla con tres columnas, la primera contiene el nombre de la propiedad, la segunda el tipo y la tercera el valor. En la tercera es donde Ud. debe registrar los valores de los metadatos. La Figura 9.4 muestra estas pestañas.



Descripción Propiedades personalizadas

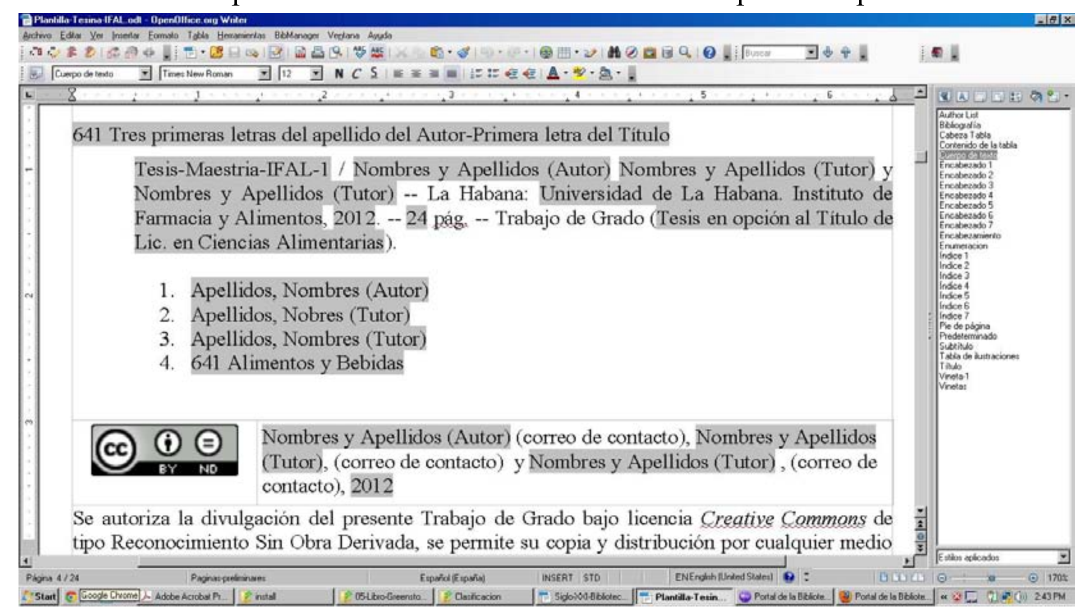


Figura 9.4 Página legal del documento de tesina. Los textos marcados con fondo gris se corresponden con los metadatos o descripciones que se registran automáticamente a partir de las propiedades del documento. En la parte superior se muestra el contenido de las pestañas: Descripción y Propiedades personalizadas.

A continuación se describen detalladamente los contenidos de los metadatos del documento.

- **Descripción del Documento.**
  - **Título**: se registra el título completo del documento.
  - **Tema**: Se registra la clasificación temática del documento, por ejemplo, “*641 Alimentos y Bebidas*”, donde 641 es la clasificación numérica del Dewey que se corresponde con la temática de los Alimentos. Para más información consulte la tabla de clasificación Dewey en la dirección: http://bives.mes.edu.cu/install/Plantilla-Tesina-IFAL.odt
  - **Palabras Clave**: mínimo de tres palabras clave que identifican el contenido del documento, por ejemplo: Tecnología de los alimentos, Nutrición, Análisis sensorial.
- **Propiedades personalizadas del documento.**
  - **Autor (es)** del trabajo de diploma en los formatos: “Nombres y apellidos” y “Apellidos,

Nombres (Autor)”
- **Tutor (es)** del trabajo de diploma en los formatos: “Nombres y apellidos” y “Apellidos, Nombres (Tutor)”
- **Universidad donde defiende** el trabajo de diploma, por ejemplo: “Universidad de La Habana. Instituto de Farmacia y Alimentos”
- **Signatura Topográfica** (Se emplea para la localización en el fondo de la biblioteca) en las universidades cubanas se emplea el Código Dewey-Tres primeras letras del autor, y primera letra significativa del título, por ejemplo, si el título y el autor son: Los alimentos funcionales … de Raul Torres, la signatura topográfica es: **641-Tor-A**
- **Año en que se defendió la tesis:** por ejemplo 2012.
- **Grado académico por el cual se opta**, por ejemplo: Tesis en opción al título de Lic. en Ciencias Alimentarias.
- **Título de la tesina truncado** a una línea (para los encabezamientos de página y el enlace de la portada.

Una vez insertado en el documento los metadatos o la descripción del documento, estos serán insertados en el cuerpo del documento donde correspondan (véase Fig. 9.1). Concluida la inserción de metadatos, se procede a la escritura o incorporación de los datos de la tesis.

9.6.1.2- Datos o contenido del documento de tesis

Los estudiantes que acostumbran a utilizar el *Microsoft Word* para la escritura de sus trabajos de diploma pueden continuar haciéndolo de esta manera, aunque es mejor hacerlo directamente sobre el LibreOffice. Pero en ocasiones no se dispone de este software o se trabaja en computadoras que no lo tienen instalado. En ese caso pueden continuar trabajando en Word y luego copiar y pegar los contenidos sobre la plantilla en LibreOffice. Esta práctica fue calificada como apropiada y amigable por la mayoría de los estudiantes encuestados durante las pruebas realizadas en la Facultad de Farmacia y Alimentos de la Universidad de La Habana (9). Sólo para el caso de las figuras y las tablas se encontraron algunas dificultades debido a la falta de conocimiento del trabajo con figuras y tablas, más que por problemas con la técnica de copiar y pegar.

La plantilla define la estructura capitular, los diferentes epígrafes preliminares y todos los datos que se requieren para un trabajo de diploma. Simplemente se debe copiar y pegar, o escribir sobre las plantilla. Uno de los aspectos más importantes se refiere a la estructura capitular y la selección de los títulos y sub-títulos, ya que a partir de estos se construye la “Tabla de Contenidos” y los “Bookmark” o puntos de acceso para la navegación dentro del documento.

A la hora de componer el documento se debe tener especial cuidado de **no borrar** los formatos de los títulos, para eso es necesario conocer su estructura. En la Tabla 9.4 se presentan los nombres y la descripción de cada uno de los estilos definidos en la plantilla para las tesinas.

Tabla 9.4 Estilos definidos en la plantilla de tesinas

Nombre del estilo	Lugar donde se utiliza
Autores	Portada y portadilla.
Bibliografía	Referencias bibliográficas
Contenido de la Tabla	Tablas
Cuerpo de texto	Texto del documento.
Encabezado 1	Portada, sólo para la construcción del primer bookmark, es invisible!

Nombre del estilo	Lugar donde se utiliza
Encabezado 2	Páginas preliminares: portadilla, opinión del tutor, página legal, pensamiento, dicatoria, agradecimientos, Tabla de Contenidos, Resumen, Introducción, Desarrollo.
Encabezado 3	Primer nivel del encabezado en cada capítulo.
Encabezado ...	Niveles subsiguientes de encabezado 4, 5, 6 ...
Encabezado 7	Máxima profundidad de los encabezados que se recomienda emplear.
Encabezamiento	Encabezamiento de las páginas de contenido.
Enumeración	Listas numeradas
Índice 1	Tabla de contenido que describe el nivel de Encabezado 1
Índice …	Tabla de contenido que describe los niveles de Encabezado subsiguientes.
Índice 7	Tabla de contenido que describe el nivel de Encabezado 7.
Pie de página	Pie de las páginas de contenido.
Predeterminado	Estilo que se asume cuando no se selecciona ninguno. No se recomienda emplear en lugar del estilo Cuerpo de texto.
Subtítulo	Portada y portadilla. Se emplea para el nombre de la institución donde se defiende la tesis, el lugar y el año.
Tabla e ilustraciones	Se emplea en el pie de las figuras y el encabezamiento de las tablas.
Título	Primera página de cada capítulo. Se emplea para el título decorativo "CAPÍTULO ..."
Viñeta-1	Segundo nivel de enumeración con viñetas.
Viñetas	Primer nivel de enumeración con viñetas.

Los estilos anteriores son suficientes para lograr las estructuras con una complicación media. Si se requieren diseños más elaborados o estilos especiales simplemente se definen a partir de la modificación de los existentes. Lo más importante es mantener los diseños de los niveles de títulos.

9.6.1.3- Creación de la tesis en formato PDF

El formato PDF se ha convertido en un estándar "de facto", incluso, las versiones 1.4 (Adobe 5) y anteriores están recogidas en normas ISO, por lo que para producir documentos en estos formatos no es indispensable la utilización del software propietario Adobe Profesional ©, el propio LibreOffice "salva" directamente en formato PDF.

En este epígrafe se definen las propiedades que se requieren establecer para la exportación de la tesina en formato PDF.

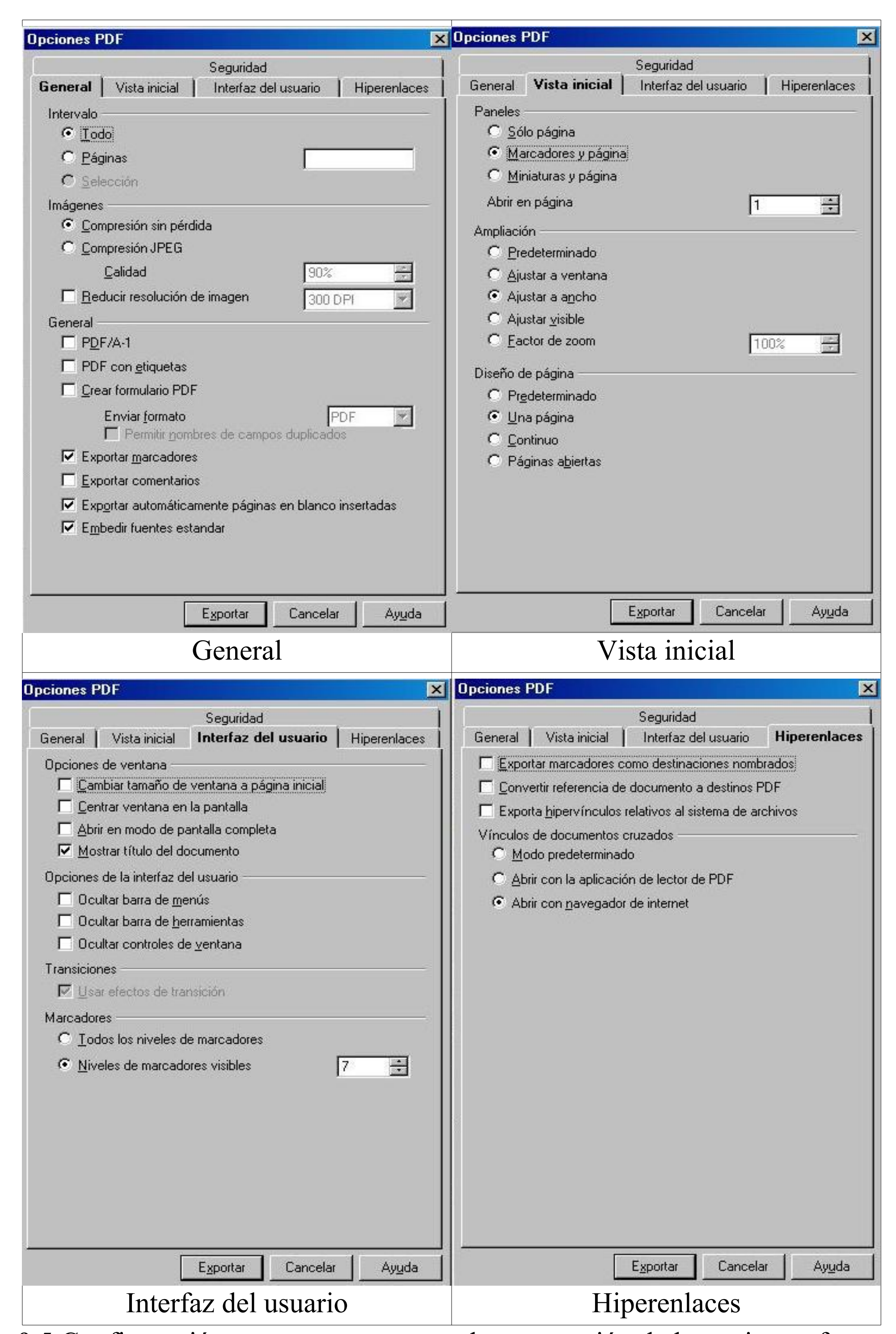
En la barra de menús buscar "Archivo" y seleccionar "Exportar en formato PDF ...", se abre una ventana con cinco pestañas:

- General,
- Vista inicial,
- Interfaz del usuario,

– Hiperenlace y
– Seguridad

En la Tabla 9.5 se presenta una vista de las cuatro primeras pestañas con la configuración que se requiere para la exportación de la tesina en formato PDF. No se requiere cambiar al configuración de la quinta pestaña, no se modifican los valores por defecto, que son: "Sin contraseña abierta configurada" y "Permisos de contraseña no configurados". De esta forma el documento no será encriptado y estará *sin restricciones de ningún tipo*. Esta configuración es indispensable para lograr que el Greenstone construya los ficheros XLM asociados a los PDF y que construya los índices por todas las palabras del texto, de gran utilidad para la realización de trabajos de inteligencia a partir de las colecciones de tesinas.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0



General | Vista inicial

Interfaz del usuario | Hiperenlaces

Figura 9.5 Configuración que se propone para la exportación de las tesinas a formato PDF

A continuación se enumeran las opciones de cada una de las pestañas de la configuración de exportación al formato PDF.

- **General**: Intervalo = Todo; Imágenes = Compresión sin pérdidas; General = Exportar marcadores, Exportar automáticamente páginas en blanco insertadas, *Embeber* fuentes (el término correcto en español es *Incrustar* las fuentes).
- **Vista inicial**: Paneles = marcadores y páginas, Ampliación = ajustar al ancho; Diseño de

página = Una página.

- **Interfaz del usuario:** Opciones de ventana = Mostrar título del documento, Marcadores = Niveles de marcadores 7.
- **Hiperenlaces:** Vínculos de documentos cruzados = Abrir con navegador de Internet.

Las tesinas exportadas en formato PDF con la configuración anteriormente descrita están prácticamente listos para ser incorporados en la biblioteca Greenstone.

Otro tipo de documento académico que se acostumbran incluir en las bibliotecas digitales universitarias es el libro de texto. Sin embargo, los autores no acostumbran publicar con la licencia de acceso abierto, pues aspiran cobrar sus derechos de copia. Esta justa aspiración en muchas ocasiones se convierte en un obstáculo para la socialización de estos documentos, sobre todo si se trata de libros de texto en formato impreso. Sin embargo, la producción de los e-libros ha comenzado a constituirse en una excelente opción, tanto para los autores como para los lectores.

A continuación se expone los procedimientos que ha desarrollado la Editorial Universitaria cubana (http://revistas.mes.edu.cu) para la producción de e-libros de texto, ya sea para su divulgación en la modalidad de acceso abierto como en acceso privativo (digital e impreso bajo demanda).

9.6.2- E-libros de texto, guías de estudio y apuntes del profesor

Romillo Tarke y cols. (10) propone un Sistema Integrado y Progresivo de Medios de Enseñanza que parte de asegurar los principales materiales en formato impreso:

> *"A los estudiantes de pregrado se les garantizará como material impreso: la guía de estudio y el texto, ambos materiales de conjunto, deben posibilitar que el estudiante pueda vencer, al nivel requerido, las asignaturas matriculadas, aun cuando no se disponga de los restantes medios antes mencionados ... En los casos de las asignaturas que requieran de la elaboración de un nuevo texto por autores cubanos, puede tomarse la decisión de fundir en un cuaderno o documento los contenidos del texto y la didáctica de la guía."*

A partir de la idea anterior se propone la siguiente modificación (véase la Figura 9.6): insertar como segundo escalón los materiales electrónicos, entre ellos el libro de texto digital, listo para ser incorporado en una biblioteca digital o para su impresión digital (también conocida como impresión bajo demanda).

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

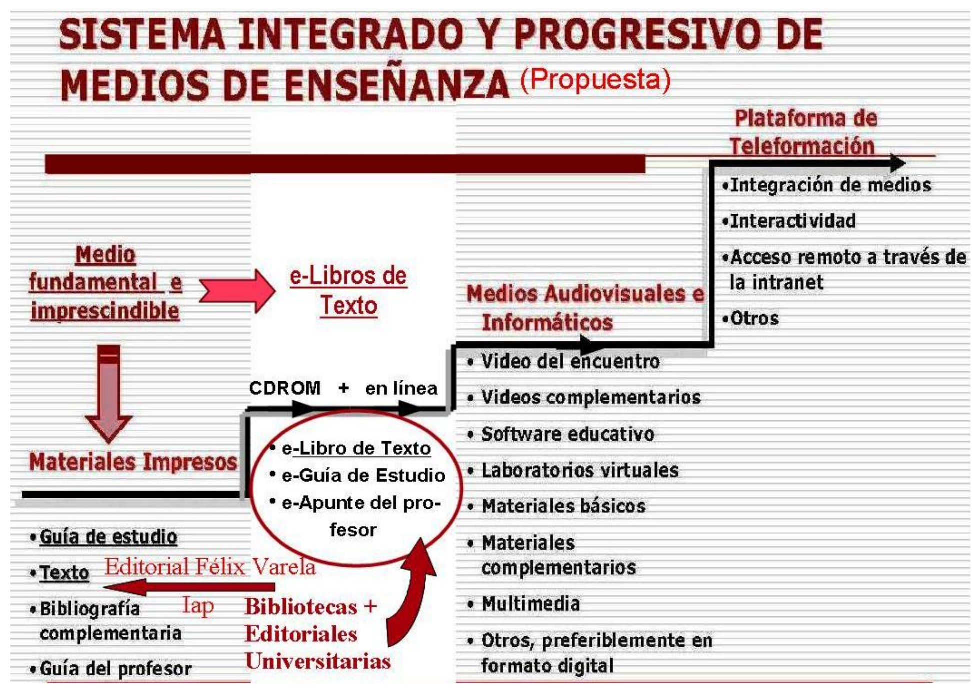


Figura 9.6 Propuesta de modificación al Sistema Integrado y Progresivo de medios de enseñanza-aprendizaje descrito por Romillo Tarke y cols. (10).

La Editorial Universitaria cubana (http://revistas.mes.edu.cu) ha desarrollado plantillas para la producción de los libros de textos universitarios. Las indicaciones para la utilización de esta plantilla son muy semejantes a las descritos para el caso de las tesinas, sólo que sus metadatos tienen algunas especificaciones adicionales, como son el pie de imprenta y el ISBN (véase la plantilla en http://bives.mes.edu.cu/install/Titulo-Libro-Texto-Plantilla-2012.odt).

Un requerimiento importante para divulgar los libros de texto en las bibliotecas digitales lo constituye el convenio que los autores firman con la Editorial que lo publica. En este caso se pone como ejemplo el convenio que la Editorial Universitaria cubana firma con sus autores para divulgar sus obras en la modalidad de acceso abierto, este convenio está en: http://bives.mes.edu.cu/install/Convenio-EDUNIV-autores.pdf).

La Editorial Universitaria incorpora soluciones técnicas para lograr la publicación de los libros académicos:

- Registro de publicaciones con e-ISBN13 en Cuba.
- Producción de PDF y ePUB con calidad comercial (aceptados por ebrary).
- Producción de registros MARC 21 con calidad comercial (aceptados por al YBP para GOBI3[9]).

La plantilla y el procedimiento desarrollado por la Editorial Universitaria cubana permite que los propios autores escriban y editen sus obras y las entreguen prácticamente listas, para su publicación con la Editorial Universitaria, como para su incorporación en bibliotecas y repositorios elaborados con el software Greenstone.

[9] GOBI3 (Global On-line Bibliographic Information) es la herramienta para la “gestión de colecciones” en el mundo académico. Disponible en: http://www.gobi3.com

9.7- Consideraciones finales

Para crear una colección de libros digitales en formato PDF mediante el software Greenstone es necesario lograr que los documentos que se pretenden incorporar cumplan con determinados requisitos de calidad para asegurar la construcción de los índices a texto completo y permitir a los usuarios que puedan copiar y pegar textos independientemente de la plataforma que empleen. En este capítulo se describe detalladamente un procedimiento, a partir del empleo tanto de software libre como propietario, que logra esto y asegura la calidad comercial de los documentos (los documentos son aceptados por la plataforma comercial e-Libro).

9.8- Referencias Bibliográficas

(1) Adama Web S.L. (2011). Difusión y divulgación científica en Internet. España: Gobierno del Principado de Asturias.

(2) Melero, Remedios (2005). Significado del acceso abierto (open access) a las publicaciones científicas: definición, recursos. El Profesional de la Información, 15 (4), 255-266

(3) Rojas V., M. Alejandra & Rivera M., Sandra (2011). Guía de buenas prácticas para revistas académicas de acceso abierto. Santiago de Chile: ONG Derechos Digitales.

(4) Torricella Morales, Raul G. y Lee Tenorio, Francisco. Repositorios de Recursos Educativos: abiertos y privativos. Ponencia presentada en el Sexto Seminario-Taller del Sistema de Bibliotecas del MININT. Centro de Eventos CELAMED, del Centro de Investigaciones Médico - Quirúgicas (CIMEQ), La Habana, 15 al 17 de Mayo 2012.

(5) Torricella Morales, Raúl G.; Lee Tenorio, Francisco; Carbonell De La Fé, Sergio (2008). Infotecnología : la cultura informacional para el trabajo en la Web. La Habana : Editorial Universitaria. 2008. -- 50 pág. – ISBN 978-959-16-0742-3

(6) Tramullas J. Biblioteca Digital Greenstone. Capitulo 10. En.Software Libre para Servicios de Información [Libro en Internet] Digital. Madrid: Pearson Prentiss Hall, 2006. Disponible en: <http://tramullas.com >

(7) Cruz Santos, Ileana Dayamina de la (2011). Propuesta de modelo de gestión de información digital agraria cubana. Instituto Superior de Tecnologías y Ciencias Aplicadas (Tesis en opción al Título de Máster en Gerencia de la Ciencia y la Innovación Tecnológica). – La Habana : Editorial Universitaria, 2011. -- ISBN 978-959-16-1366-0. -- 134 pág.

(8) ABBYY FineReader 8.0 Profesional Edition. Sistema de reconocimiento óptico de caracteres. (c) 2005 ABBYY Software, reservados todos los derechos.

(9) Torricella Morales, Raúl Gonzalo; Hernández Monzón, Aldo y Huerta Espinosa, Victor Manuel (2009). Propuesta de procedimiento para la digitalización de los trabajos de diploma: caso de estudio. La Habana : Editorial Universitaria, ISBN 978-959-16-1034-8.

(10) Romillo Tarke, Antoni, Batista T., Nuri & Gonzáles Fernández-Larrea, Mercedes (2008). El perfeccionamiento de los procesos formativos en las Sedes Universitarias Municipales con el empleo de las tecnologías educativas. Palacio de las Convenciones: Congreso Universidad 2008.

(11) Díaz Mayans, Concepcion y Torricella Morales, Raul Gonzalo (1996). Las nuevas tecnologías de la información en la educación superior. La Habana: Editorial Universitaria, 1996.

(12) Fernández Fernández, Maikel Manuel, Torricella Morales, Raul G. y Abreu Bartomeo, Yanedi (2010). Biblioteca Digital Alma Mater: subproyecto Conceptualización del Portal de

Aplicaciones Educativas y la Biblioteca Digital Alma Mater. Marco referencial (conceptualización para antecedentes) para las soluciones a entregar. Caracas (Venezuela): ALBET Ingeniería y Sistemas.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

CAPÍTULO 10

10- Búsqueda de Información en Greenstone

Lic. Elvia Vázquez Velázquez (elvia@cecam.sld.cu) Licenciada en Letras. Centro de Cibernética Aplicada a la Medicina (CECAM) I.S.C.M.H. Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354

Palabras clave: Bibliotecas Digitales, Colecciones de documentos, Greenstone, Búsqueda y recuperación de información

10.1- Introducción

El término búsqueda de información, difundido con inestimable importancia cada vez más por estos días, incluye el proceso de recuperar datos con exactitud en lo encontrado con gran utilidad para satisfacer las necesidades de información del usuario en documentos que contienen una información determinada; constituye a su vez una excelente vía para reconocer la elevada cantidad de información existente en el mundo en formato digital así como identificar, localizar fuentes de información relevantes y utilizar de forma efectiva las tecnologías de información para tal empeño.

Los sistemas de recuperación de información solicitan que el usuario introduzca palabras clave o términos significativos relacionados con el tema que el programa buscará en los registros de la base de datos. La búsqueda será efectiva en la medida en que las palabras proporcionadas permitan recuperar exclusivamente la información que se necesita a una gran velocidad, y presentar los registros en que fueron encontradas. Es muy importante seleccionar el campo en el que se buscará, se puede buscar en varios simultáneamente ellos pueden ser: título, autor, tema, fecha, descripción, editoriales, etc. Esto permite especificar más la búsqueda.

La búsqueda de información se efectúa mediante sistemas de búsqueda informativa. El sistema Greenstone, considerado uno de los más completos para la creación y presentación de colecciones digitales, facilita a los usuarios varias maneras de encontrar información. Se puede tener acceso a la mayoría de las colecciones mediante las funciones de búsqueda y consulta (1).

Los textos de los documentos incluidos en las colecciones están destinados para recuperar la información deseada y en Greenstone se le conoce como “Términos de búsqueda”.

Este programa resulta idóneo para atesorar miles de documentos a texto completo; por eso es indispensable conocer opciones y recursos que ofrece para obtener provecho de este en cuanto al aspecto de acceder y recuperar, de forma eficiente, la información almacenada en sus colecciones, pues cuenta con clasificadores y motores de búsqueda indispensables para obtener este fin.

Greenstone por defecto presenta una interfaz en la que se pueden observar botones, formularios, una barra de navegación, un texto que describe e identifica a la colección y modos posibles para desarrollar búsquedas, esta interfaz es semejante para todas las colecciones que incluyen la biblioteca.

La barra de desplazamiento que aparece en la parte superior de la pantalla en las colecciones

incluye la función ***búsqueda,*** siempre está disponible, seguida de botones convenientes a todos los clasificadores que se han definido.

Encontramos en el extremo derecho superior varios botones también presentes en todas las páginas ellos son: ***Principal, Ayuda Preferencias*** Ver Figura 10.1. La función de estos es proporcionar el acceso a páginas que se corresponden con estos calificativos.

PRINCIPAL AYUDA PREFERENCIAS

Figura 10.1. Botones disponibles en todas las páginas de la interfaz.

Botón Principal. Enlaza a la página principal de Greenstone, a su vez permite seleccionar desde aquí cualquier otra colección presentes en la biblioteca digital.

Botón Ayuda. Con este se puede acceder a un texto de ayuda descrito en esta colección, informa cómo se pueden leer los documentos, o cómo proceder a efectuar la búsqueda de determinadas palabras entre otros temas que pueden orientar y esclarecer posibles dudas del usuario.

Botón Preferencias. Este botón posibilita establecer algunos cambios cuantas veces lo desee con diferentes opciones en la interfaz de usuario, además de personalizar la búsqueda según intereses del beneficiario.

En todas las páginas puede visualizar una barra navegación de color verde, en la figura del ejemplo se observan 6 botones: el primero corresponde a la función ***Buscar*** y los otros 5 enlazan con los 5 botones de navegación determinados para esta colección durante el proceso de configuración de la colección, los que facilitan la exploración de las listas de documentos. (Véase Figura 10.2).Esta interfaz mostrada en la figura puede variar de una colección a otra según criterios de selección y configuración de clasificadores.



Figura 10.2 Botones o clasificadores y función Búsqueda en barra de navegación en la colección.

Reconozcamos ahora los botones de la barra verde.

Botón Búsqueda(siempre presente).Enlaza con la página de búsqueda; admite elaborar búsquedas por palabras específicas comprendidas en el texto de un documento proporcionado, ya sea por frases o por párrafos.

Botón Titulo(añadido) Expone una lista de los títulos de documentos organizados por orden alfabético, responden con suficiente precisión al contenido de las páginas, por tanto la recuperación de la información tendrá mayor relevancia y precisión.

Botón Autores (añadido) .Con un clic sobre este botón se visualiza una lista de autoría de cualquier tipo de documento, que puede ser individual como corporativa, organizados también por orden alfabético.

Botón Tema (añadido).Proporciona la lista de temas o palabras clave referidas a las temáticas de los diferentes textos que incluyen la colección; estas son importantes pues son una de las formas más idóneas y protegidas de dar a conocer el contenido de cualquier tipo de documento y se produce una recuperación exitosa con un valioso índice de precisión de la información.

Botón Fecha (añadido). Facilita la lista de documentos por fechas de publicación, esta resulta siempre de mucha importancia para los usuarios de dicha información, pues manifiesta el grado de

caducidad o actualidad de los documentos, además de constituir un campo de recuperación de la información por la fecha de publicación.

En este caso se añadió también el botón ***fuente internet*** para incorporar el URL de donde se obtuvieron los documentos PDF descargados de la Web.

A todos ellos se accede con un solo clic.

La información ***Acerca de esta colección, (***este texto puede sustituirse por otro en el ejemplo se ha modificado por ***Descripción de esta colección***) se antepone a la descripción de la colección en la que se describe el contenido de la materia incluida en la colección.

Posteriormente aparece otro subtítulo o encabezamiento: ***¿Cómo encontrar información?***, esta sección describe brevemente las 6 maneras ventajosas para encontrar información en la colección, en la que hace referencia a los botones mencionados anteriormente, como lo presenta la figura 4. Recuerde que la cantidad depende de la cantidad de botones configurados por usted y que aparecen en la barra verde descrita anteriormente.

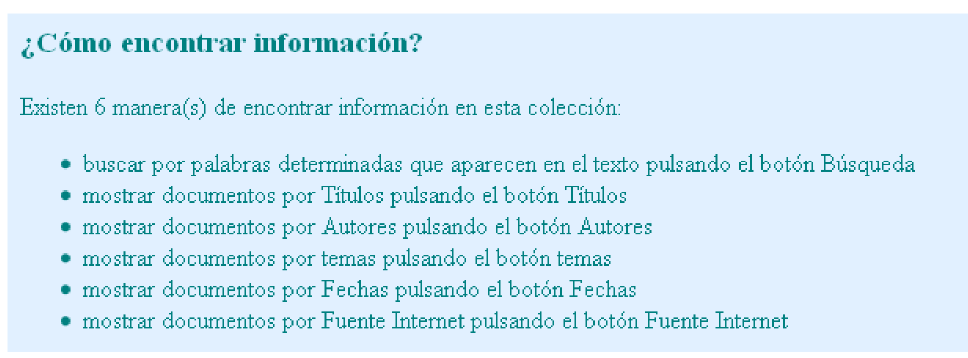


Figura 10.3 Explica la manera de encontrar información en la colección

Anteriormente explicamos la barra de navegación de color verde presentes en todas las páginas generadas por Greenstone con las diferentes opciones por las cuales resulta posible buscar en una colección. Enfatizamos que con el ejemplo que proponemos los botones utilizados se nombran: ***Búsqueda***, ***Título***, ***Autores, Tema, Fecha, Fuente internet*** es necesario resaltar que estas opciones las configura Ud. al crear la colección y que resultan significativas, porque permiten recuperar la información de forma específica y oportuna además de posibilitar la navegación de las listas jerárquicas generadas según los criterios establecidos. Estas opciones dependen de los clasificadores y los índices de búsqueda seleccionados durante el diseño y creación, internamente ligados a los metadatos asignados.

El botón ***Buscar*** Buscar en texto Este botón sirve para realizar la búsqueda pensada por el usuario en el ***Cuadro de texto***. (Figura 10.4)

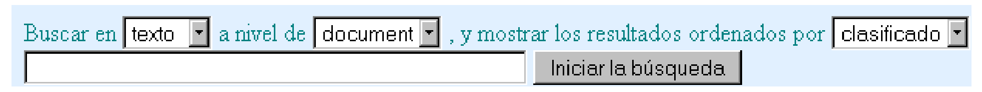


Figura 10.4 Barra de navegación Opción Buscar

En este espacio en blanco el usuario escribe los términos o palabras clave para declarar la búsqueda; pueden ser palabras simples o frases completas. (Figura 10.5)



Figura 10.5 Botón Buscar con los posibles índices de búsqueda

Como puede distinguir la figura anterior muestra los índices de búsqueda configurados durante la creación de la colección; pero existen otras alternativas para buscar por títulos de secciones o por capítulos de libros, etc, recuerde que para esto hay que examinar otros metadatos desde el panel Diseño, opción ***índice de búsqueda***.

Las búsquedas pueden realizarse de forma sencilla o avanzada. Las búsquedas sencillas se formulan con la adición de uno o varios caracteres alfabéticos y numéricos y se omiten por ejemplo signos de puntuación.

En el Botón ***Ayuda*** mencionado anteriormente expone que pueden desarrollarse 2 tipos de búsqueda o consulta:

Consultas basadas en todas las palabras: El programa busca los documentos que contengan todas las palabras especificadas en el texto y recupera los documentos que poseen esas palabras, se presentan por orden de construcción.

Consulta basada en algunas palabras. Busca en los textos si algunos de las palabras aparecen en los documentos en que se está indagando. En este caso los documentos recuperados se presentan según el orden en que se encuentran con relación a la consulta efectuada, los documentos aparecerán en un orden que corresponde a su grado de relevancia con respecto a la consulta.

Al formalizar la búsqueda de información a través de los diferentes botones y opciones mencionados, el usuario puede encontrar distintos íconos que le permiten acceder de forma factible a los diversos documentos, es recomendable aquí darle a conocer la identificación de alguno de ellos.

Puede encontrar botones que permiten depurar más la búsqueda ellos son los que representa la figura 6 siguiente

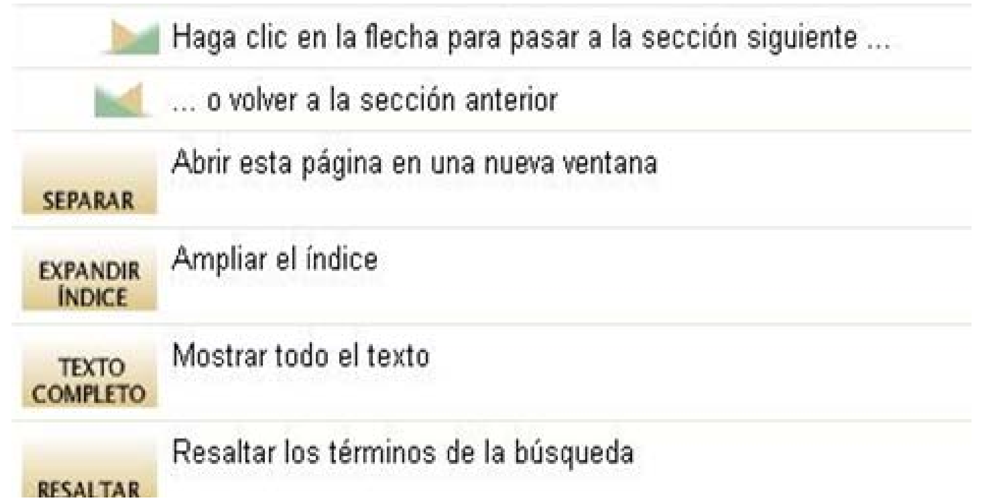


Figura 10.6 Algunos botones encontrados en el proceso de búsqueda de información.

Los términos a utilizar pueden ser todos los términos de búsqueda que considere conveniente.

El sistema Greenstone ofrece la posibilidad de búsqueda avanzada para perfeccionar la recuperación de información y trazar estrategias precisas a la hora de la búsqueda. Se activan desde el botón ***Preferencias*** y están planteadas de acuerdo con el diseño de la colección.

Para fijarlas de un clic en el botón ***Preferencias***

PREFERENCIAS

situado en la parte superior derecha de la página y pulsar en el botón ***Fijar preferencias***,

fijar preferencias

(Ver figura 7), a continuación activar o desactivar las casillas disponibles en la interfaz según su agrado. Después pulsar en cualquiera de los restantes botones de la barra de navegación. Incluye opciones referidas al modo y estilo de búsqueda, reconocimiento de mayúsculas y minúsculas, terminaciones de palabras, repara en la acentuación o no de los términos, si van a ser ignorados o no, historial de búsquedas ejecutadas y presentación de los resultados en cuanto a número de documentos y de páginas. Ver figura 7



Figura 10.7 Interfaz para realizar búsqueda avanzada con botón Preferencias

Una aclaración pertinente es que no debe nunca usar el botón *Atrás,* de la página sencillamente anula las preferencias que acaba de designar.

10.2- Indizadores MG, MGPP y Lucene

Greenstone admite la búsqueda de texto, por todo el cuerpo del documento de la colección, puede ser desde una ventana de navegación Web u optar entre los diversos índices constituidos a partir de diferentes porciones de los documentos.

En la parte superior derecha de la pantalla "***índices de búsqueda***" dentro de la pestaña ***Diseño*** aparecen los indizadores que tiene en ese momento la colección, ellos son: ***MG***, ***MGPP*** y ***Lucene***. Tal como se ven en la figura 10.

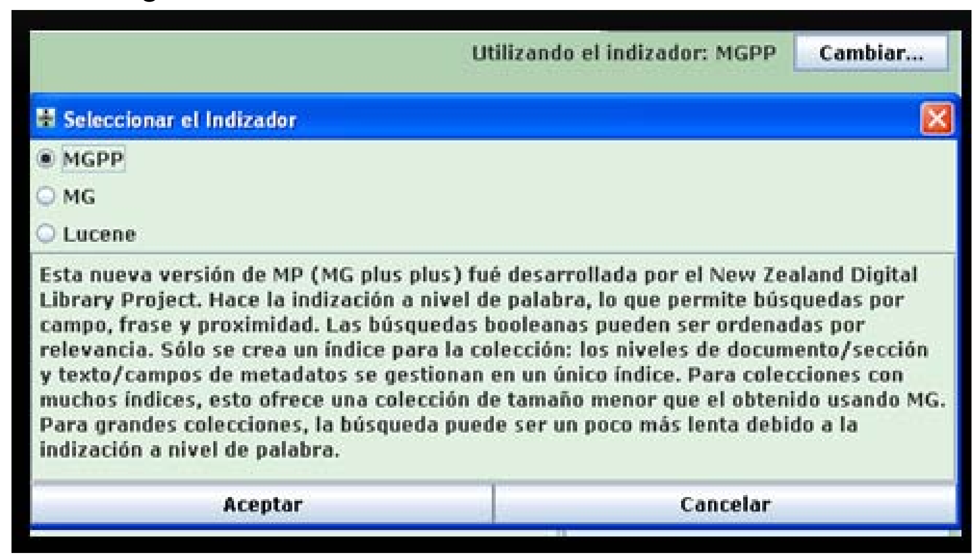


Figura 10.8 Indizadores asignados por Greenstone

***MG*: indizador original de Greenstone**

MGPP: proporciona índices a nivel de palabras y permite la proximidad, la frase, cuando efectúa la búsqueda de información. Ignora diferencias entre minúsculas y mayúsculas. Indizador por defecto para las nuevas colecciones.

Lucene: Funciona similar en la búsqueda a MGPP con la adición de comodines de un sólo carácter y por rangos dentro del texto. Facilita la incrementación de colecciones que los anteriores no lo permiten.

Cambiarlos, afecta la forma en que se construyen los índices y también las funcionalidades de búsqueda.

Según seleccione uno de ellos así será la forma en que se construyen los índices, esto se traduce en la página de preferencias de la colección, en cuanto a: idiomas, el uso de acentos, de mayúsculas y minúsculas, entre otras.

Lucene (Ver figura 9) es un indizador muy avanzado y tiene muchas ventajas; supera a ***MG*** y ***MGPP*** pues tiene muchas más opciones de búsqueda, es muy rápido y es incremental.

Independientemente de estas ventajas hay que tener en cuenta que cuando se opte por este, solamente se puede usar como servidor el Apache y que si desea exportar una colección a un CD/DVD no funcionará con él, para ello es mejor elegir el indizador ***MGPP***.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

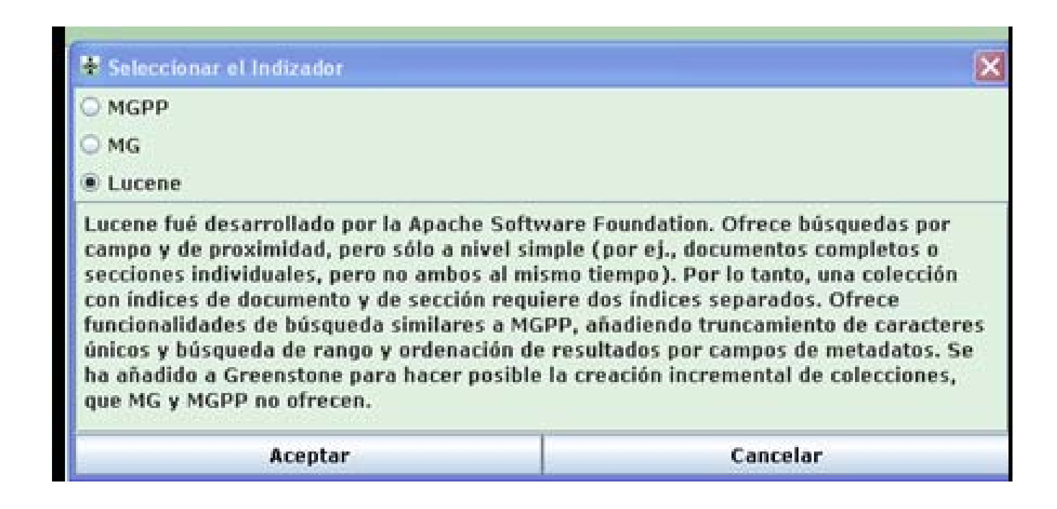


Figura 10.9 .Indizador Lucene

En la parte inferior de esta pestaña (*indices de búsqueda*) aparecen las casillas de *Opciones de indización* o *Indexing options* que afectan a todo el conjunto de índices *Cadena o Stem*: palabras por aproximación, solamente funciona en Inglés y Francés. *May/Min* o *Casefold:* Discriminación de mayúsculas. *Coincidencia de Acentos* o *Accent fold*: discriminación de acentos. *Segmentación de texto CJK* o *CJK Segmentation*

Y ***Niveles de indización*** en el que puede seleccionar: ***Documento*** o ***Sección***. Puede elegir uno u otro como predefinido tal como aparece en la fig. 10 y 11.



Figura 10.10 Presenta casillas de *Indexing options (inglés)*

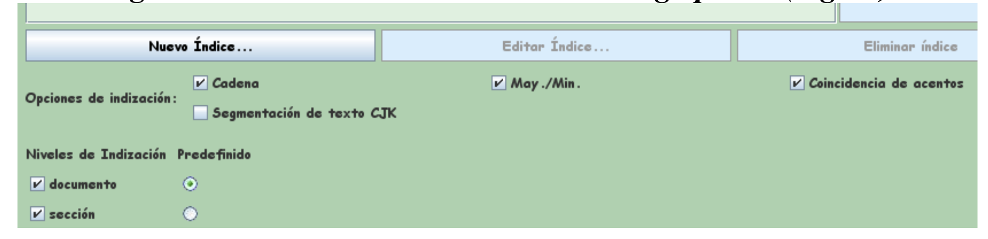


Fig 11 Opciones de indización (español)

10.3- Término de búsqueda y tipos de búsqueda

La opción ***Índices de Búsqueda o Consulta*** presentes en la pestaña Diseño de GLI conforma la parte de los documentos que Greenstone dispone para que el usuario pueda recuperar información posteriormente hasta en volúmenes extensos. Especifican criterios que se emplearán para la búsqueda y están estrechamente relacionados con los metadatos de acuerdo con las necesidades de

los usuarios.

Inicialmente Greenstone tiene tres ***índices de consultas o búsqueda*** (Véase figura 14): ***texto, títle y Source***, *índices* que contemplan las búsquedas o consultas a texto completo, por el metadato título y por el metadato fuente, el índice de texto no debe eliminarlo nunca pues determina las búsquedas a texto completo y debe situarse en primer lugar siempre en la relación de índices asignados.

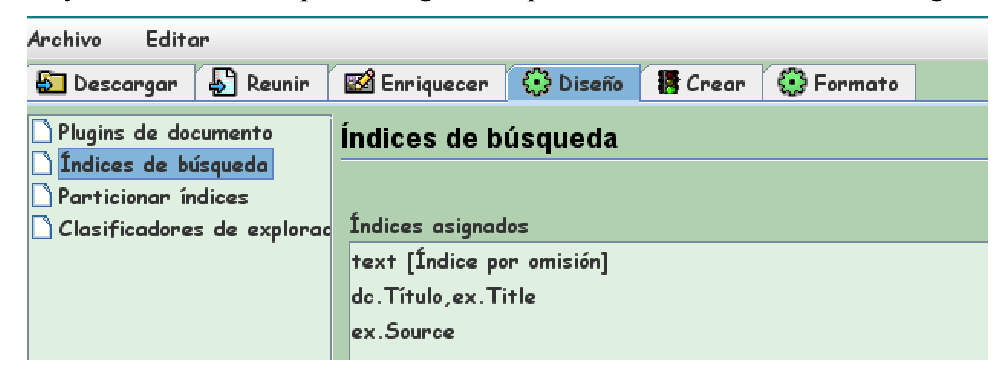


Figura 10.12. Disposición de los índices de búsqueda originales

Pueden configurarse índices que permitan consultar por varios campos a la vez y por textos completos en todos los documentos de una colección; Greenstone proporciona esto en todas las colecciones. Pueden modificarse los índices originales y sustituir por índices según gustos y necesidades de recuperación de información, para el ejercicio se han modificado y nombrados título, autor, palabras clave y fecha tal y como puede verse en la Figura 10.13 que a continuación se muestra.

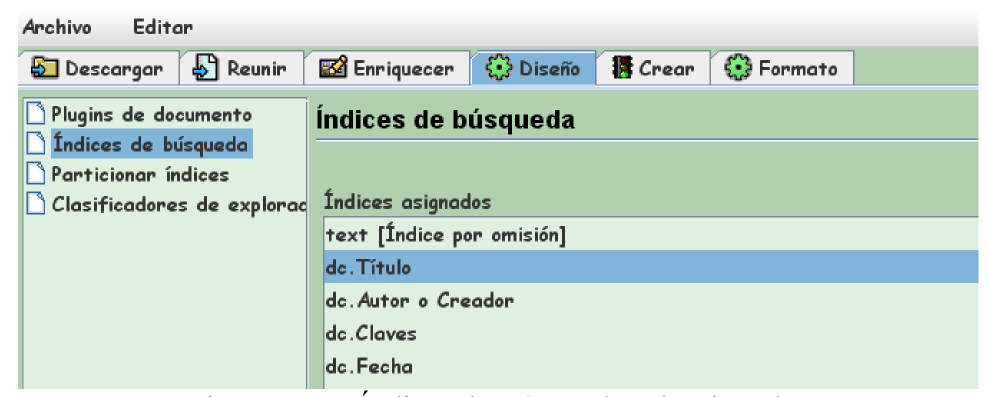


Figura 10.13 Índices de Búsqueda seleccionados

El orden de los clasificadores en la barra de navegación de la colección se define aquí y está en dependencia del gusto del creador de la colección.

10.4- Clasificadores de exploración

La opción "***Clasificadores de exploración***" sirve para asignar "clasificadores", son los que sirven para crear los índices de consulta de la colección, y se usan para navegar y buscar en ella.

Son botones para mostrar los documentos de la colección al usuario, según los criterios de ordenación predefinidos.

Los clasificadores ofrecen una ordenación y clasificación de todos los documentos. Se utilizan durante la creación de la colección organizándolos por título, autores, palabras clave, fechas, fuente…, etc, u otro metadato determinado durante la configuración y diseño de la colección; de ahí

la importancia de estos últimos en la configuración de los clasificadores.

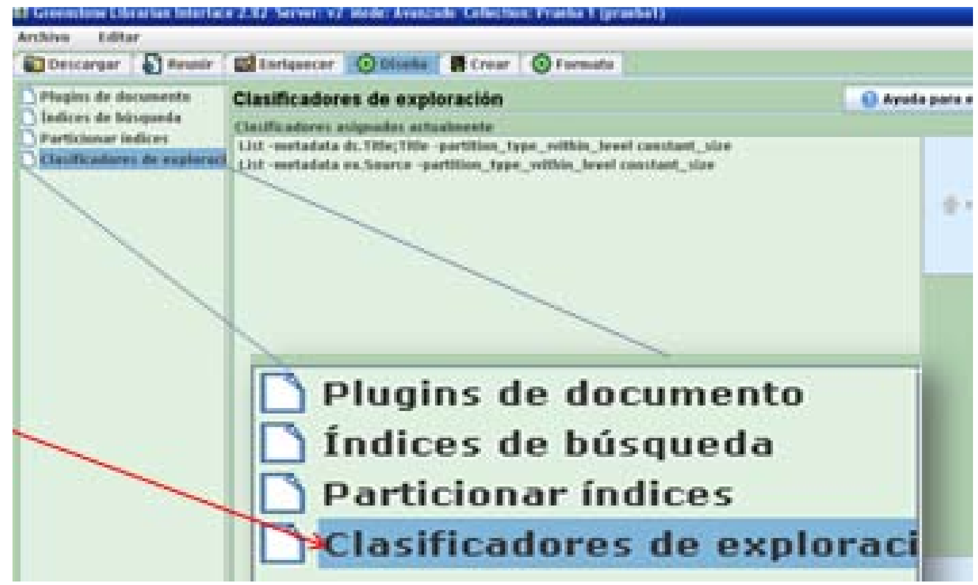


Figura 10.14 Opción Clasificadores de exploración

Para quitar uno de la lista, selecciónelo y haga clic en "*Eliminar el clasificador seleccionado*". Ver figura 15

Si desea añadir un clasificador nuevo, elíjalo por medio de la lista que ofrece la pantalla en el extremo inferior derecho y haga clic en "***Añadir el clasificador especificado***". Ver figura 15

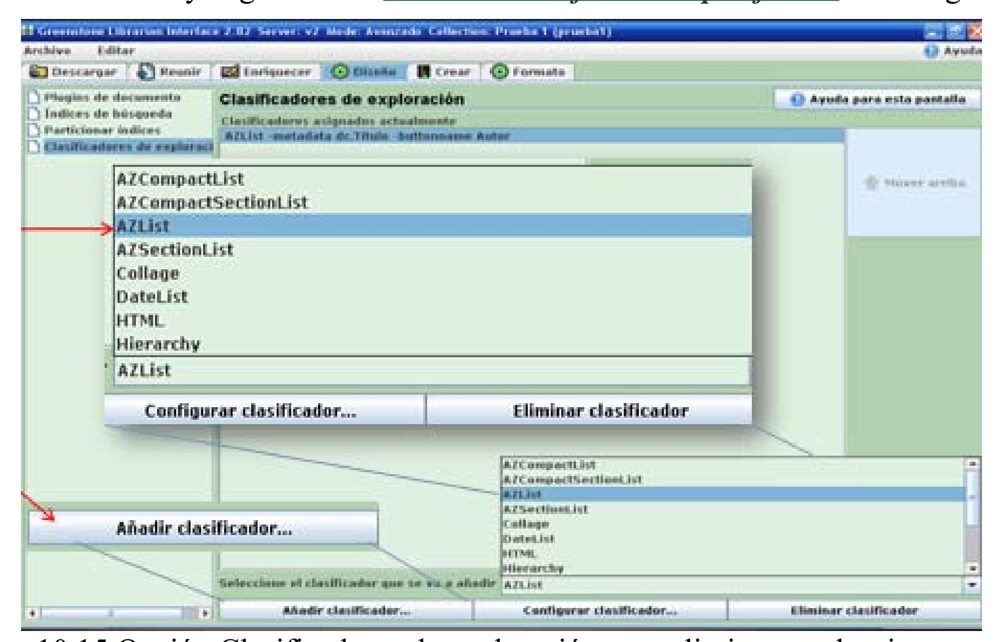


Figura 10.15 Opción Clasificadores de exploración para eliminar y seleccionar nuevos.

Algunos tipos de clasificadores que brinda Greesntone

- AZList: lista de documentos dividida por intervalos alfabéticos
- AZCompactList: lista ordenada de la A a la Z, agrupa según número de documentos para metadato autor o creador, palabras clave
- List: lista alfabética simple de los documentos
- DateList: lista ordenada por fecha

- Hierachy: lista jerárquicas, por orden de metadatos
- … Section … : usados cuando se trata de documentos con secciones y subsecciones

Existe además un tipo especial de clasificador llamado "***Phind***" (Figura. 16), que crea un índice de frases a partir de una palabra seleccionada.



Figura 10.16 Clasificador Phind , para índices de frases.

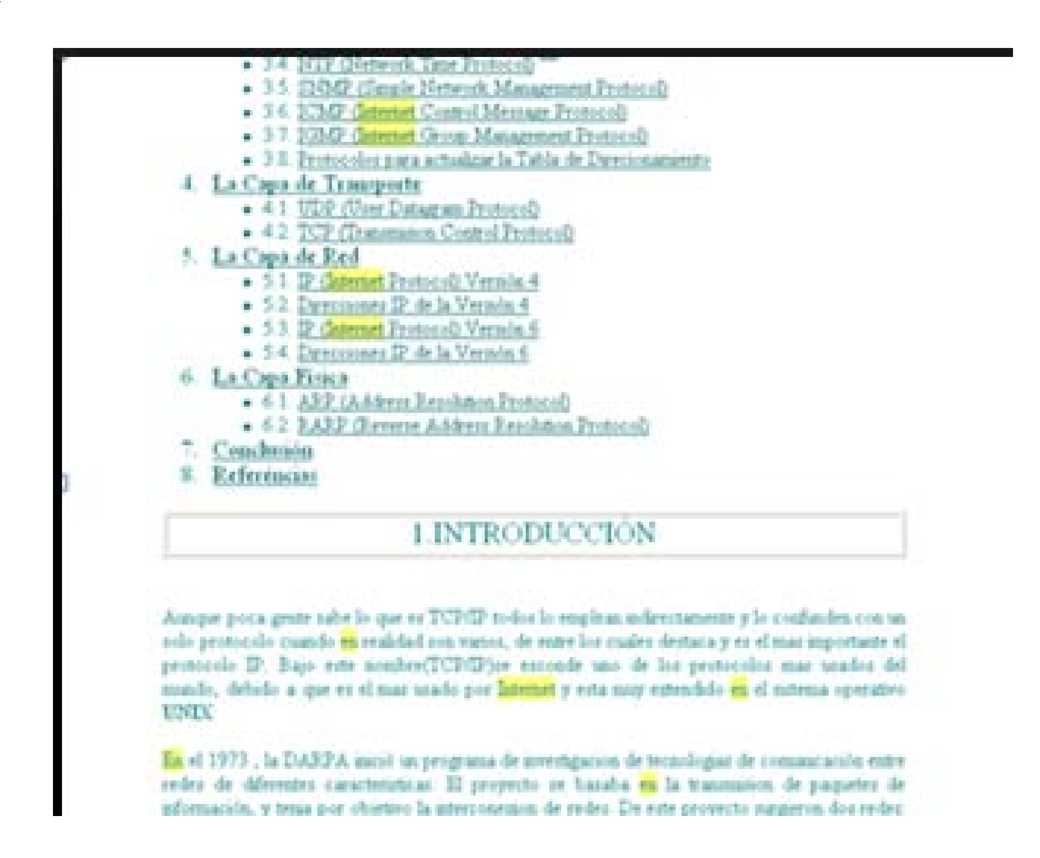
Si da un clic en Buscar aparecerá una lista de frases con el término indicado en la parte inferior como se observa en la Figura 10.17.



Figura 10.17.Lista de frases con clasificador Phind

La frase se puede extender más y más. Al final aparece una lista de los documentos que contiene dicha frase en color azul como se ve en la figura anterior (17), puede extender más esta lista si da un clic en los documentos restantes.

La frase o palabra solicitada aparece en varias partes del texto de los documentos como se ve señalado en color amarillo en la Figura 10.18,en la que se ha efectuado la búsqueda con la palabra "*Internet*" y esto es lo que muestra al usuario, tanto en subtítulos, como en el cuerpo del texto del documento.


- 3.4 NTP (Network Time Protocol)
- 3.5 SNMP (Simple Network Management Protocol)
- 3.6 ICMP (Internet Control Message Protocol)
- 3.7 IGMP (Internet Group Management Protocol)
- 3.8 Protocolos para actualizar la Tabla de Direccionamiento

4. La Capa de Transporte
- 4.1 UDP (User Datagram Protocol)
- 4.2 TCP (Transmission Control Protocol)

5. La Capa de Red
- 5.1 IP (Internet Protocol) Versión 4
- 5.2 Direcciones IP de la Versión 4
- 5.3 IP (Internet Protocol) Versión 6
- 5.4 Direcciones IP de la Versión 6

6. La Capa Física
- 6.1 ARP (Address Resolution Protocol)
- 6.2 RARP (Reverse Address Resolution Protocol)

7. Conclusión
8. Referencias

1.INTRODUCCIÓN

Aunque poca gente sabe lo que es TCP/IP todos lo emplean indirectamente y lo confunden con un solo protocolo cuando en realidad son varios, de entre los cuales destaca y es el mas importante el protocolo IP. Bajo este nombre(TCP/IP)se esconde uno de los protocolos mas usados del mundo, debido a que es el mas usado por Internet y esta muy extendido en el sistema operativo UNIX.

En el 1973 , la DARPA inicio un programa de investigacion de tecnologias de comunicación entre redes de diferentes caracteristicas. El proyecto se basaba en la transmision de paquetes de informacion, y tenia por objetivo la interconexion de redes. De este proyecto surgieron dos redes

Figura 10.18 Visualización de búsqueda con la palabra "internet"

Todos los clasificadores generan una estructura jerárquica para mostrar un índice de consulta en las diferentes colecciones.

10.5- Ejercicios de búsqueda en Greenstone según metadatos y clasificadores

Los términos de búsqueda a utilizar son palabras clave indizadas y que están incluidas en el metadato ***palabras clave***. Utilizar estas palabras puede realzar la calidad del proceso de recuperación de información.

En el Botón ***Búsqueda*** de la barra de navegación tecleamos la palabra *Metodología* y el resultado lo muestra la figura 19 siguiente



Figura 10.19 Búsqueda de información simple botón Búsqueda

El resultado de la consulta realizada relacionó a 40 documentos incluidos en la colección con el descriptor solicitado y fue efectuada con el índice de búsqueda ***texto.***

Sin embargo ese mismo descriptor fue utilizado para el índice de búsqueda ***título*** y la información muestra 8 documentos como resultado, tal y como lo presenta la figura 20.

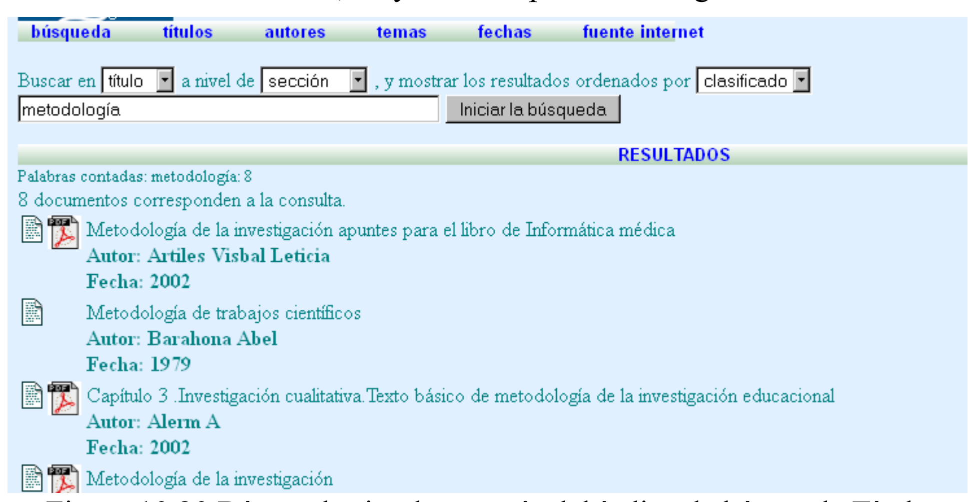


Figura 10.20 Búsqueda simple a través del índice de búsqueda Título

Hemos visualizado sólo 4 en pantalla _ por problemas de espacio_ con la intención de acentuar la información de 8 documentos del total incluidos en la colección.

Si continua la búsqueda con la misma sintaxis ***metodología*** en el índice de búsqueda temas el resultado es el siguiente.



Figura 10.21 .Búsqueda con índice tema

De todas formas se estableció la recuperación de información solamente por los índices de búsqueda, puede hacerlo también por los diferentes botones (clasificadores) visualizados en la barra de búsqueda.

10.6- Formulario Avanzado de Búsqueda

Si en cambio queremos ajustar la interfaz para realizar una búsqueda avanzada hemos fijado las preferencias con el botón ***Preferencias*** de la siguiente manera

Preferencias de búsqueda

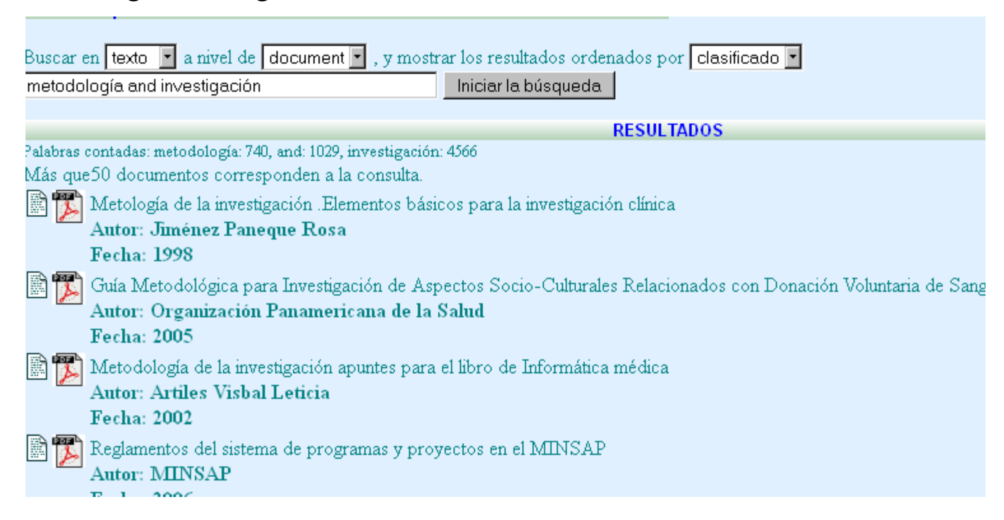
Modo de búsqueda:	◯ modo de búsqueda simple ◉ búsqueda avanzada (permite efectuar búsquedas booleanas utilizando los caracteres !, &, | y los paréntesis)
Estilo de búsqueda:	◉ normal con caja query ◯ con campos fields
Reconocimiento de mayúsculas y minúsculas:	◉ ignorar diferencias entre mayúsculas y minúsculas ◯ deben corresponder mayúsculas y minúsculas
Terminaciones de las palabras:	◯ ignorar las terminaciones de las palabras ◉ debe coincidir la palabra completa
Diferencias de acentos;	◯ ignorar acentos ◉ las tildes deben coincidir
Historial de búsquedas:	◉ no mostrar el historial de búsquedas ◯ mostrar registros del histórico de búsquedas

Presentar hasta 50 resultados con 20 resultados por página.

Realizamos la búsqueda con los siguientes descriptores:

metodología ,***investigación*** y se utiliza además el operador **AND**.

El resultado fue el siguiente Figura 10.22


Buscar en texto a nivel de document , y mostrar los resultados ordenados por clasificado
metodología and investigación Iniciar la búsqueda
RESULTADOS
Palabras contadas: metodología: 740, and: 1029, investigación: 4566
Más que50 documentos corresponden a la consulta.
Metología de la investigación .Elementos básicos para la investigación clínica
Autor: Jiménez Paneque Rosa
Fecha: 1998
Guía Metodológica para Investigación de Aspectos Socio-Culturales Relacionados con Donación Voluntaria de Sang
Autor: Organización Panamericana de la Salud
Fecha: 2005
Metodología de la investigación apuntes para el libro de Informática médica
Autor: Artiles Visbal Leticia
Fecha: 2002
Reglamentos del sistema de programas y proyectos en el MINSAP
Autor: MINSAP

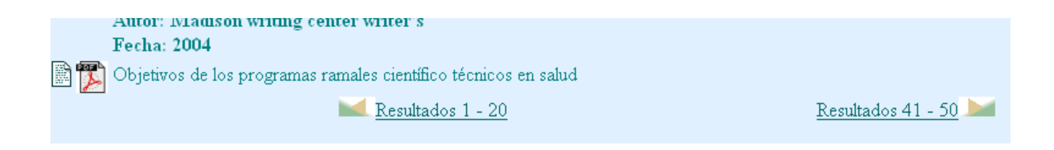


Figura 10.22 Búsqueda avanzada con el índice de búsqueda Texto

Lo que implica mayor cantidad de documentos como resultado de la consulta realizada. En la parte inferior presenta los resultados divididos en una primera página hasta 20

Resultados 1 - 20

que a su vez nos enlaza a otra página que llevaría desde 41 hasta 50 documentos,

Resultados 41 - 50

como lo presenta la Figura 10.21.

Como se puede comprobar las opciones de búsqueda en Greenstone son variadas solo debe establecer la manera de efectuar la búsqueda de información de la forma que le resulte más conveniente y apropiada para su investigación.

Consideraciones finales

Diversas opciones le permiten al usuario recuperar información de las colecciones creadas con Greenstone ya sea si establece formularios de búsqueda durante el proceso de configuración y creación de estas o a través de los diferentes campos de los índices de búsqueda de la colección con determinada configuración establecida desde el botón Preferencias. Las posibilidades ofrecidas por este programa con los mecanismos de búsqueda resultan eficientes para la recuperación de información para el usuario.

10.7- Referencias bibliográficas

(1) Ian H. Witten y Stefan Boddie Biblioteca digital Greenstone .En Guía del usuario. New Zealand : University of Waikato. Digital Library. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Develop_es_index.html >

(2) Bainbridge D, McKay D, Witten H. I. Guía del programador. Biblioteca digital Greenstone. [Sitio Web en Internet].Departamento de Informática. Universidad de Waikato, Nueva Zelanda. Marzo de 2004.Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Develop_es_all.html >

(3) Ejercicios de tutoría de Greenstone (Marzo de 2006) .Disponible en: <http://greenstonewiki.cs.waikato.ac.nz/wiki/gsdoc/GNUFDL.html >

(4) Biblioteca digital Greenstone. Guía del programador Chapter 2 Explotación óptima de sus documentos. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_getting_the_most_out_of_your_documents.htm>

SECCIÓN II: LAS FORMAS AVANZADAS DE TRABAJO CON GREENSTONE

"Si consigo ver mas lejos es
porque he conseguido auparme a
hombros de gigantes"

Isaac Newton (1642-1727) Matemático y Físico británico

http://www.proverbia.net/citasautor.asp?autor=707

CAPÍTULO 11

11- Publicaciones electrónicas y derechos de autor y de propiedad

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH)

Palabras clave: *Copyright, Copyleft, Creative Commons*, Derecho de Autor, Desarrollo Sostenible, Economía Digital, Estrategias de Inclusión Digital, Publicaciones Electrónicas , Sociedad de la Información.

11.1- Introducción

Puede considerarse que existen varios tipos de publicaciones electrónicas (1): páginas Web, archivos de ayuda (con extensión .hlp), Directorios, Revistas y otras.

En 1989, Brailstord emitió el siguiente concepto abarcador de lo que pudiera representar una publicación electrónica" trasciende la "*oficina sin papeles*" o la pura transferencia y diseminación electrónica de información mediante redes, e incluye todos los métodos asociados a la producción de documentos con el uso de computadoras" (2).

11.2- Publicaciones electrónicas, criterios de clasificación. Ventajas y desventajas

11.2.1- Criterios de Clasificación de las Publicaciones Electrónicas

Pudiera hablarse de una clasificación de las publicaciones electrónicas relacionada con la forma, contenido, soporte y presentación de la información (3), como detallaremos a continuación.

Por su forma: estaría determinado por el formato electrónico en que está almacenada la publicación y el software browser necesario para visualizarla correctamente.

Por su contenido: estaría determinada por la esencia de su contenido, así como por la periodicidad y la seriedad de la edición. Las publicaciones pueden ser formales, informales, revistas científicas o de vulgarización, libros, periódicos, preprints, obras de referencia en línea, etcétera.

Por su soporte: está determinado por el medio físico que almacena la información: un servidor de servicios de Internet, disquetes o CD-ROM (*Compact Disk Read-Only Memory*).

También pudiéramos considerar la propuesta de Hechavarria (4) de clasificación de las publicaciones electrónicas atendiendo a:

- Forma de acceso / recuperación / distribución
- Forma de almacenamiento
- Formato

11.2.2- Ventajas y desventajas de las publicaciones electrónicas

Resultan muy sucintas las consideraciones hechas por Hechavarria (4) al respecto y por ello las describimos aquí:

11.2.2.1- Ventajas

- Oportunidad y rapidez en la difusión y distribución de la información.
- Inclusión de información tridimensional, sonido y video, así como la facilidad de su manipulación.

11.2.2.2- Desventajas

- Es indispensable un equipo para su lectura.
- No existe una reglamentación, se carece de un código de ética establecido.
- Favorece la proliferación de la información, además de que esta puede ser modificada por los usuarios.
- Se dificulta la identificación entre la información primaria y la secundaria.

Pueden ser también consideradas como ventajas (5,6):

- Realizar búsquedas en el texto completo.
- Acceder instantáneamente, sin necesidad de desplazamiento.
- Vincular referencias, citas y publicaciones.
- Enlazar recursos relacionados como bases de datos u otros materiales complementarios como películas y animaciones, que facilitan la expresión de ideas difíciles de plasmar en un formato impreso.
- Relacionar autores y lectores por correo electrónico, se favorece entonces la comunicación científica.
- Publicar inmediatamente, a partir de un régimen de edición continua.
- Realizar correcciones y comentarios, seguir las ideas y sugerencias hechas por los lectores.
- Disminuir los costos y el consumo de papel, se hacen copias impresas sólo de los artículos que realmente son de interés.

11.3- Derecho de autor y derecho de copia .Las iniciativas internacionales sobre los derechos. Declaración de la Interdependencia Digital

En nuestros días hablamos de "derecho de autor" y queda definido como el conjunto de normas y principios que regulan los llamados derechos patrimoniales que la ley concede a los autores; por el solo hecho de la creación de una obra ya sea literaria, artística, o científica, tanto publicada o que todavía no se haya publicado (7).

Los primeros antecedentes que reporta la historia relacionado con este concepto se remontan al siglo XVIII en Inglaterra cuando en aquel entonces, los "libreros" se adjudicaban el derecho a perpetuidad de controlar la copia de los libros que habían adquirido los autores (8).

Es necesario destacar que, la protección internacional del derecho de autor es automática, nace en el momento de la creación de la obra y este principio se aplica a todos los países como Parte del Convenio de Berna (9).

Resulta curioso el hecho de que el propósito original de lo que se denominó en la constitución de Estados Unidos, copyright o parte patrimonial del derecho de autor, fue establecido como una

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

iniciativa para promover el progreso e incentivar a los autores hacia la publicación de sus resultados (10).

El copyright ha sido hasta nuestros días, una especie de monopolio temporal amparado por las Constituciones de los diferentes países, que cuando expira permite que la obra pueda ser utilizada en el dominio público, el cual se encuentra exento de regulaciones. (11).

De acuerdo con la American Library Association (ALA), la nueva era digital presenta nuevos retos con relación a las doctrinas fundamentales del denominado copyright, que son obstáculos a los servicios bibliotecarios (12), que son recogidas en el Digital Millenium Copyright Act (13).

¿Que obras son protegidas por el denominado copyright?. El copyright protege los trabajos originales de los autores (12), que están expresados de forma tangible. Esto incluye las siguientes categorías:

- Trabajos literarios
- Trabajos musicales, incluyendo las palabras acompañantes
- Trabajos dramáticos incluyendo la música acompañante
- Trabajos de pantomima y coreografía
- Trabajos pictóricos, gráficos y esculturales
- Películas y otros trabajos audiovisuales
- Grabaciones de música
- Trabajos arquitectónicos

¿Que obras no son protegidas por el denominado copyright?. Esto incluye las siguientes categorías (12):

- Trabajos que no tienen una expresión tangible (por ejemplo trabajos coreográficos que no han sido anotados, descritos o grabados)
- Títulos, nombres, frases cortas, slogans, símbolos familiares o dibujos
- Ideas, procedimientos, métodos, sistemas procesos, conceptos, principios, descubrimientos o dispositivos que no sean distinguidos por una descripción, explicación o ilustración.
- Trabajos que consisten solo en información , que es una propiedad común y no presentan una autoría (por ejemplo, calendarios estándar, cartas de peso y talla, listas o tablas tomadas de documentos públicos o de otras fuentes comunes.

Se considera que existen varias instituciones vinculadas de un modo u otro a proteger o bien a tratar de legislar los atributos del derecho de autor (14): la OMPI u Organización Mundial de la Propiedad Intelectual, con sede en Ginebra y conformada a su vez por 16 organismos especializados del sistema de las naciones Unidas; la UNESCO, intentado conservar el patrimonio de cada nación y la difusión del conocimiento; la IFLA (Federación internacional de Asociaciones de Bibliotecas e Instituciones) potenciando el uso libre de la información en los servicios bibliotecarios; la ALA (American Library Association) con su importante iniciativa sobre la Ley sobre el Derecho de Copia en el Milenio Digital (The Digital Millenium Copyright Act) que entro en vigor en el año 2000 en Estados Unidos y otras iniciativas mas. Si quisiera obtener una información mas detallada sobre este interesante tema lo invitamos a consultar el trabajo original de Morales E, García J.F. (14).

11.4- El "*fair use*" o "uso ético" de la información. *Copyleft*

Los tres más importantes derechos públicos son: el dominio público, el uso razonable (o "*fair use*") y el denominado derecho de compra (11).

La existencia del copyright solo confiere al autor de la obra la paternidad y el derecho de decidir las modalidades de uso de la misma, y da potestad de modo singular a los "libreros" de nuestros tiempos, de gobernar los designios de una obra, hasta que se agote este derecho y la obra pase al dominio publico.

Esta situación en una economía digital en la que se pretende que la información y el conocimiento sean patrimonio de todos, resulta altamente contradictoria. Y existen criterios de algunos especialistas en el tema que consideran que resulta imposible mantener los derechos de autor tal y como fueron construidos al comienzo del capitalismo industrial (15, 16).

Debemos considerar que los derechos de propiedad según el sentido exacto de lo que el término representa, comprenden (7):

- todas las formas de delimitación de los usos (usus), del disfrute del producto y de la alienabilidad de los activos;
- las condiciones de aplicación de las reglas y las normas que determinan el usus, el fructus y el abusus de todo tipo de bien (entendido en un sentido lato como aquello que es objeto de una valorización simbólica, social y económica).

Y considerar además que todos estos aspectos son legislados en cada país de acuerdo a arreglos particulares lo cual dificulta aun mas en una sociedad que pretende compartir el conocimiento, llegar a un consenso de cómo efectuar este proceso.

En la obra "Capitalismo Cognitivo, propiedad intelectual y creación colectiva", Blondeau (17) se refiere a la teoría de Harold Demsetz, (quien fuera uno de los grandes iniciadores neoclásicos de una refundición de la teoría de la propiedad), quien considera a su vez, que la propiedad no es mas que «la libertad de ejercer una elección sobre un bien o un servicio», para destacar que en el capitalismo actual, el carácter absoluto y totalitario, de esta libertad sobre la propiedad; debe recaer sin ningún obstáculo sobre las tres dimensiones de un bien o de un servicio, y se refieren entonces a: el uso, el fruto que se puede obtener directamente o por delegación y la cesión total o condicional.

Esta renovadora idea sobre el concepto y uso de la propiedad han sido plasmadas evidentemente en las categorías "fair use" y copyleft.

Ampliaremos sobre el concepto de "fair use" cuando abordemos en un siguiente punto la iniciativa Creative Commons, aunque entendemos que ambas categorías están vinculadas, ya que en el nuevo concepto de copyleft emitido y defendido por los propugnadores de software libre, encabezado por Richard Stallman y el proyecto GNU que dirige (18), consideramos, se juega fundamentalmente con la idea de "uso razonable" cuando se formula esta nueva concepción.

De acuerdo a las palabras del propio Stallman, el Copyleft (18) es un método para convertir un programa en software libre y exigir que todas las versiones del mismo, modificadas o ampliadas, también lo sean. Copyleft [`o `izquierdo de copia"traducido literalmente] es la regla que implica que, cuando se redistribuya el programa, no se puedan agregar restricciones para denegar a otras personas las libertades centrales. Esta regla no entra en conflicto con las libertades centrales, sino que más bien las protege.

Si analizamos el concepto de copyleft, nos percatamos que también presenta una consonancia con el denominado "derecho de compra", el cual establece que cualquiera que obtenga la copia de un trabajo puede hacer lo que estime conveniente con el, aunque el suministrador se encuentre en contra de ello (11),

Y resulta curioso que el concepto de copyleft confiere plenas posibilidades a la persona que obtiene

una obra bajo estas circunstancias, para modificarlo o reestructurarlo como desee.

No queremos terminar esta sección sin mencionar que ha sido creada la denominada Fundación Copylef (19) que pretende potenciar y defender la producción de arte, cultura y ciencia bajo licencias copyleft.

11.5- "Creative Commons" o Bienes Comunes Creativos

¿Que representa Creative Commons?



La visión de esta entidad es desarrollar todo el potencial de Internet. Es decir acceso universal a la educación, la investigación y, participación plena en la cultura, enfocada a una nueva era de desarrollo crecimiento y productividad (20).

La misión, desarrollar, soportar y proporcionar la estructura legal y técnica que maximice la creatividad, la innovación y uso compartido en la era digital (20).

¿De que modo puede lograrse esto?

La infraestructura que provee *Creative Commons* se basa en un grupo de licencias y herramientas sobre el copyright, que crean un balance dentro del tradicional "all rights reserved" que han establecido las leyes sobre copyright creadas (21)

¿En que principios se basa esta infraestructura?

- Dominio Público
- Uso Ético
- Derechos Morales
- Derechos de Publicidad

Podemos encontrar la definición de estos principios en el propio sitio Web de esta institución, pero detallémoslos aquí (21):

- **Dominio Público**: un trabajo se encuentra en el dominio público cuando es libre para su uso por cualquiera y para cualquier propósito sin restricción de copyright.
- **Uso Ético**: todas las jurisdicciones que persiguen algún uso limitado de los materiales con copyright, sin permiso. La licencia Creative Commons, no afecta los derechos de los usuarios bajo las limitaciones del copyright y sus excepciones, tales como el *fair use* (uso ético) y el *fair dealing*; que en las cortes se relaciona con el numero y el impacto económico que representa el uso de algún material que presenta derecho de autor, cuando no es significativo el impacto, se dice entonces que es un uso o fair dealing (22).
- **Derechos Morales**: además de los derechos de aquellos que establecen las licencias para obtener la eliminación del nombre de un trabajo, cuando es usada de forma derivada o colectiva en forma no deseada, las leyes de copyright en la mayoría de las jurisdicciones alrededor del mundo (con la notable excepción de Estados Unidos, excepto en muy pocas circunstancias limitadas) el premio a los creadores "Derechos Morales" puede proveer alguna zona limite , si un trabajo derivado representa un "derogatory treatment" (relacionado con el fenómeno de mutilación , alteración o distorsión de un trabajo que pueda perjudicar la reputación del autor (23).
- **Derechos de Publicidad**: estos desprotegen a los individuos sobre el control de su voz, imagen o semejantes usados para fines comerciales. Con Creative Commons será necesario obtener un permiso para hacer uso en estos casos.

¿Todo en Creative Commons resulta beneficioso?

Sánchez (24) relaciona los siguientes elementos como adversos:

- esta licencia no es una respuesta alternativa al copyright, ya que se sustenta en él para delimitar los derechos que pueden, o no, tener terceros sobre la obra creada
- esta licencia aún considera los bienes culturales como mercancías que son necesarias para favorecer una producción creativa, situación que excluye obras producidas por culturas ajenas a este concepto, como algunas asiáticas, africanas e indígenas

11.6- Las iniciativas internacionales para compartir información (Repositorios Digitales, Open Access ,OCLC, WorldCat, Protocolo OAI, Protocolo Z39.50)

11.6.1- Repositorios digitales

En algunas disciplinas, principalmente en ciencias, han surgido los llamados archivos o repositorios de documentos electrónicos como alternativa para la rápida comunicación de resultados de investigación (25). Esos documentos se han llamado eprints de forma genérica. Este nuevo concepto agrupa tanto aquellos documentos que no han pasado por un proceso de certificación o peer review (preprints) como aquellos que si han pasado por tal proceso o postprints (artículos, libros, etc.).

Un repositorio, depósito o archivo es un sitio Web centralizado donde se almacena y mantiene información digital, habitualmente bases de datos o archivos informáticos. Pueden contener los archivos en su servidor o referenciar desde su Web al alojamiento originario (26).

Un usuario que busca información en un repositorio, describe de manera breve el documento que quiere obtener como respuesta. El resultado es que el usuario "navega" en el repositorio realizando consultas y refinando resultados hasta que queda satisfecho (27).

- OPENDOAR es un Buscador de Repositorios, disponible en: http://www.opendoar.org/
- e-PRINTS es una Recopilación de Repositorios de e-Prints, disponible en: http://www.eprints.org/
- OAIster Es un proyecto financiado por la fundación Mellon con el objetivo de crear una amplia colección de recursos digitales gratuitos, útiles y que previamente eran de muy difícil acceso y ponerla al alcance de cualquier usuario de la forma más sencilla posible. Es decir, trata de sacar a la luz colecciones que antes eran invisibles. Todos los recursos tienen el texto completo disponible en la red de forma que siempre se pueda llegar a los contenidos. Recoge datos de todos los archivos conocidos. En total 122 archivos con más de un millón de registros. Disponible en: http://oaister.umdl.umich.edu/

Se difunde la iniciativa de creación de repositorios institucionales, como un intento más de facilitar el acceso a la información en las instituciones académicas. Estos son definidos como " una base de datos compuesta de un grupo de servicios destinados a capturar, almacenar, ordenar, preservar y redistribuir la documentación académica de la Universidad en formato digital" (28). Cuando se dice que los Repositorios Institucionales son abiertos e interactivos esto quiere decir que cumplen con el protocolo OAI (del cual hablaremos posteriormente) y permiten acceso abierto a la documentación académica (28).

11.6.2- Iniciativa Open Acces

Este movimiento promueve la publicación “abierta” de información científica a través de Internet

(29)

Directorios Open Access más conocidos (29)

- Directory of Open Access Journals (DOAJ), http://www.doaj.org/
- SCIELO, http://www.scielo.org/
- PUBMED CENTRAL, http://www.pubmedcentral.nih.gov/
- BIOMED CENTRAL, http://www.biomedcentral.com/
- AERA (Open Access Journals in the Field of Education, http://aera-cr.asu.edu/
- OCLC (Online-Computer-Library-Center)
- OCLC es una organización sin fines de lucro, que funciona como una asociación dedicada a la prestación de servicios bibliotecarios computarizados y a la investigación, con el propósito de promover el acceso del público en general a la información de cualquier lugar del mundo y de reducir los costos asociados (30).

Más de 72.000 bibliotecas en 170 países y territorios de todo el mundo han utilizado los servicios de OCLC para ubicar, adquirir, catalogar, prestar y preservar material bibliotecario (http://www.oclc.org/americalatina/es/about/default.htm)

11.6.3- WorldCat

WorldCat (http://www.worldcat.org) es el catálogo mundial del consorcio de bibliotecas OCLC, integrado por 72.000 bibliotecas en 172 países. WorldCat es una red mundial de servicios y contenidos que utiliza el Web y que pretende lograr mayor apertura y productividad en los servicios. Se considera la red más grande del mundo en contenido y servicios de bibliotecas (31).

Ventajas de WorldCat (32):

- Permite la búsqueda en varias bibliotecas a un tiempo y localiza la búsqueda en la biblioteca mas cercana
- Encuentra de igual modo libros, música y videos.
- Localiza artículos de investigación e incluso audiobooks.
- Presenta un vinculo a la sección "pregunte a bibliotecario" y a otros servicios de su biblioteca
- Puede enviarle un a revisión sobre un tema o contribuir con una información factual para Ud.

Este es el formulario de búsqueda que ofrece WorldCat.

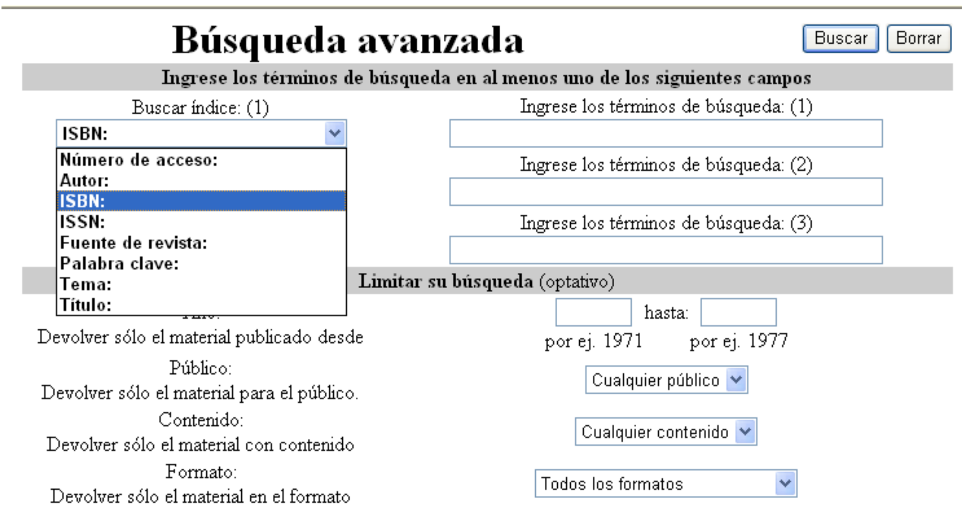


Figura 11.1.- Búsqueda avanzada [WorldCat.org].

Disponible en: http://www.worldcat.org/wcpa/rel20111201/js/autocomplete/autocomplete.j

11.6.4- Protocolo OAI (Open Archives Initiative)

La Open Archives Initiative (OAI) se creó con la misión de desarrollar y promover estándares de interoperabilidad para facilitar la difusión eficiente de contenidos en Internet (25). Surgió como un esfuerzo para mejorar el acceso a archivos de publicaciones electrónicas (eprints), en definitiva, para incrementar la disponibilidad de las publicaciones científicas. Los trabajos iniciales se centraron en el desarrollo de marcos de interoperabilidad para la federación de archivos de eprints, pronto pareció evidente que dichos marcos (permitir el intercambio de múltiples formatos bibliográficos entre distintas máquinas utilizando un protocolo común) tenían aplicaciones más allá de esta comunidad. Por ello se adoptó un objetivo mucho más amplio: abrir el acceso a un rango de materiales digitales

11.6.5- El origen de Metadata Harvesting

En Octubre de 1999 se organizó una reunión en Santa Fe (Nuevo México, USA), con la idea de que la interoperabilidad de estos archivos de eprints era clave para aumentar su impacto entre la comunidad académica. Con ella se podrían federar varios archivos, intercambiar registros o realizar búsquedas en disciplinas relacionadas al mismo tiempo. Los participantes en esta reunión fueron especialistas en bibliotecas digitales, así como representantes de los principales archivos existentes (33):

El resultado de la reunión fue un conjunto de acuerdos técnicos y organizativos conocidos como la Convención de Santa Fe. Los aspectos técnicos incluían tres puntos fundamentales: un formato para los metadatos, un protocolo basado en el antiguo Dients y un sistema de identificación (33). La solución adoptada fue la recolección de metadatos (metadata harvesting).

Básicamente OAI-PMH utiliza transacciones HTTP para emitir preguntas y obtener respuestas entre un servidor o archivo y un cliente o servicio recolector de metadatos. El segundo puede pedir al primero que le envíe metadatos según determinados criterios como por ejemplo la fecha de creación de los datos. En respuesta el primero devuelve un conjunto de registros en formato XML, incluyendo identificadores (URLs por ejemplo) de los objetos descritos en cada registro (34).

Las peticiones se emiten utilizando los métodos GET o POST del protocolo HTTP y constan de una lista de opciones con la forma de pares del tipo: clave=valor. Existen seis peticiones que un cliente puede realizar a un servidor: OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting). http://www.openarchives.org

11.6.6- Protocolo Z39.50

Z39.50 es un estándar internacional para la comunicación entre computadoras , bibliotecas y sistemas de información relacionados (35).

Fue propuesto originalmente en 1984 con el objetivo de proporcionar una vía estándar de interrogación a las bases de datos bibliográficas es mantenido por la Agencia Z39.50, que es administrada por la Biblioteca del congreso de Estados Unidos (35).

Z39.50 es una norma para consultar catálogos de bibliotecas en Internet utilizando las mismas reglas para todos.los catálogos, algo así como usar siempre la misma pantalla para consultar cualquier biblioteca. Para consultar sistemas Z39.50 es necesario contar con un programa especial llamado cliente Z39.50 o navegador Z39.50, o bien a través de páginas o de portales en Internet. Es recomendable contar con un programa de consulta porque no todas las bibliotecas son accesibles a

través de WWW (36).

Los protocolos como el Z39.50 tienen una funcionalidad muy completa, por ejemplo, tratan cuestiones como el manejo de sesiones, gestión de conjuntos de resultados y permiten la especificación de predicados para filtrar los resultados obtenidos. Sin embargo, esta funcionalidad acarrea un incremento en la complejidad de la implementación y, en consecuencia, de los costes (25)

El protocolo Z39.50 especifica el formato y los procedimientos que gobiernan el intercambio de mensajes entre un cliente y un servidor. El cliente puede enviar una búsqueda, indicar una o más bases de datos e incluir una consulta y también parámetros, los que determinan si los registros identificados por la búsqueda podrían devolverse como parte de la respuesta. El servidor responde con el número de registros identificados. El cliente puede, entonces, recuperar los registros seleccionados. El cliente asume que los registros seleccionados forman un conjunto de resultados y los registros pueden referirse por su posición dentro del conjunto (36).

Existen tres puntos importantes de las aplicaciones con este protocolo que han cambiado la actividad bibliotecaria (37):

- 1. Los clientes más modernos Z39.50 pueden enviar solicitudes a varias bibliotecas simultáneamente, sea dentro de una misma solicitud o en varias independientes. Esto permite un ahorro de tiempo.
- 2. El formato básico utilizado para el intercambio de registros bibliográficos es el MARC. La capacidad de presentar y transferir en formato MARC, permite al cliente utilizar esa información para un procesamiento posterior.
- 3. Con el uso de Z39.50 se han abierto otros procesos bibliotecarios, particularmente, la conmutación bibliográfica y el préstamo bibliotecario.

11.6.7- Consideraciones Finales

En la construcción de una Sociedad de la Información /Conocimiento donde resulta esencial asegurar la creación del conocimiento, compartir y transferir este conocimiento, adquiere gran trascendencia, y es un reto, la transformación de las vigentes reglas sobre los derechos de autor, que actualmente representan un freno al pleno disfrute y aprovechamiento del caudal de conocimiento humano. Las iniciativas internacionales que favorecen el uso de la información deben contribuir a la implementación de estrategias de inclusión digital que favorezcan el desarrollo de la sociedad que avizoramos.

11.7- Referencias Bibliográficas

(1) Travieso Aguiar M. Las publicaciones electrónicas: una revolución en el siglo XXI. ACIMED [Seriada en Internet] vol. 11No.2-3. 05, Enero 2003. Disponible en: <http://www.bvs.sld.cu/revistas/aci/vol11_2_03/aci010203.htm >

(2) Brailstord, D.I. y R.I. Beach. Electronic publishing. A journal and its production. Computer Journal 32 (6):482-493, Dec. 1989.

(3) Pérez Mariño J. M. Actualidad y perspectivas de las publicaciones electrónicas. ACIMED [Seriada en Internet] vol. 6 No.2, 1998 Disponible en: <http://bvs.sld.cu/revistas/aci/vol6_2_98/aci05298.htm>

(4) Hechavarría Kindelán A. Una propuesta de concepto y clasificación de publicaciones electrónicas. Cienc. Inform. 1997; 28 (2):36-43.

(5) Ordoñez C.A. Edición de Publicaciones Digitales. Beneficios y desventajas. [Libro en Internet] UNAM, 2005, pág. 49-52. Disponible en: <http://www.mati.unam.mx/index.php?option=com_content&task=view&id=204&Itemid=122>

(6) Sauval M. Internet, una revolución tecnológica en el campo de las publicaciones electrónicas [Sitio Web en Internet]. 1998. Disponible en: <http://psiconet.com/michael/congreso/genova98.html >

(7) Boutang Y. Nuevas tecnologías de la información y la comunicación. Enfoque económico. Boletín del Instituto de Investigaciones Económicas. [Seriada en Internet].11 de febrero de 2004. Disponible en: <http://www.iiec.unam.mx/notiiiec/extras/conferencia_Yann_Moulier_Boutang.pdf>

(8) Wikipedia Foundation. Derecho de autor y Copyright. [Enciclopedia en Internet]. 28 may 2006. Disponible en: <http://es.wikipedia.org/wiki/Derecho_de_autor>.

(9) Berne Convention. Berne Convention for the Protection of Literary and Artistic Works. [Website in Internet]. International Bureau of WIPO. September 9, 1886, (amended on September 28, 1979). Available from: <http://www.wipo.int/index.html>.

(10) Stallman R. E-Books Freedom Or Copyright [Website in Internet]. GNU. 2005/04/26. Available from: <http://www.gnu.org/philosophy/ebooks.html >.

(11) Besser H. Intellectual Property: The Attack on Public Space in Cyberspace. [Website in Internet]. UCLA School of Education & Information. 3/19/01. Available from: <http://www.gseis.ucla.edu/~howard/Copyright/>.

(12) ALA. Libraries & Copyright. What Is Copyright? [Website in Internet]. 2011. Available from: <http://www.ilovelibraries.org/ill.xml 2011>.

(13) ALA. American Libraries - Digital Copyright Bill Passed by Congress [Website in Internet]. October 19, 1998. Available from: <http://www.ala.org/Template.cfm?Section=alafaq&template=/cfapps/faq/faq.cfm >.

(14) Morales E. García J. F.La propiedad intelectual y los derechos de autor: iniciativas de organismos y asociaciones internacionales. Infodiversidad [Seriada en Internet].Numero 010. 2006. Disponible en: <http://redalyc.uaemex.mx>.

(15) Boutang Y. Riqueza, propiedad, libertad y renta en el capitalismo cognitivo en: Capitalismo cognitivo, propiedad intelectual y creación colectiva. [Libro en Internet]Edición: Traficantes de Sueños Madrid; Abril de 2004. Disponible en: <http://traficantes.net7>.

(16) Torres M. El libro y los derechos de autor en la Sociedad de la Información. Novedades del Programa CTS+I. [Sitio Web en Internet]. . Organización de Estados Iberoamericanos Presentado en: I Reunión de Ministros Iberoamericanos de Sociedad de la Información. Madrid, España, 27 y 28 de septiembre de 2001. Disponible en: <http://www.campus-oei.org/salactsi/mtorres.htm>.

(17) Blondeau O. Génesis y subversión del capitalismo informacional. [Libro en Internet]. Capitulo 1. En: Capitalismo cognitivo, propiedad intelectual y creación colectiva. Edición: Traficantes de Sueños (1ª edición).Madrid. Abril de 2004. Disponible en: <http://traficantes.net>

(18) Stallman M. ¿Qué es el copyleft? [Libro en Internet].Capitulo 14 En: Software libre para una sociedad libre. Versión 1.0 Primera edición en castellano (papel): Edición electrónica a cargo de: Miquel Vidal. Edición: Traficantes de Sueños Madrid. Noviembre 2004. Disponible en: <http://sindominio.net/traficantes>.

(19) Fundacioncopyleft.org. Fundación Copyleft [Sitio Web en Internet].2010. Disponible en:<http://fundacioncopyleft.org/es>

(20) Creative Commons.org. . About - Creative Commons [Website in Internet]. 01/03/2011. Available from: <http://creativecommons.org/wp-content/themes/cc6/style.css?>

(21) Creative Commons.org. Creative Commons License Deed. Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0). [Website in Internet]. 1999/02/22. Available from: < http://creativecommons.org/licenses/by-nc-nd/3.0/>.

(22) Intellectual Property Office - Permitted uses of copyright works - fair dealing. [Website in Internet]. 2006-09-011 Available from: < http://www.ipo.gov.uk/copyright-whatsnew-rss.xml>.

(23) Commonwealth Consolidated Acts. COPYRIGHT ACT 1968 - SECT 195 A J. Derogatory treatment of artistic work [Website in Internet]1.1.1968. Available from: <http://www.austlii.edu.au/au/legis/cth/consol_act//ca1968133/>.

(24) Sanchez M. " Creative Commons". Desventajas. [Sitio Web en Internet]. Grupo de investigación Copia/Sur .2008. Disponible en: <http://observatoriotecedu.uned.ac.cr/index.php/actualidad/propiedad-intelectual/170-creative-commons.html>.

(25) Creative Commons.org. Repositorios digitales. [Sitio Web en Internet]. 2008. Disponible en: <http://creativecommons.org/licenses/by-nc/3.0/ >.

(26) Vieyra G.E. Repositorios de Información [Sitio Web en Internet]. 2000-05-23. Disponible en: <http://www.fismat.umich.mx/~elizalde/tesis/node7.html >

(27) Barrueco J.M. Subirats I. OAI-PMH: Protocolo para la transmisión de contenidos en Internet

(28) Barton M.R., Waters M.M. Cómo crear un Repositorio Institucional. Manual LEADIRS II[Libro en Internet].2004. Disponible en:< http://www.mit.edu/libraries>

(29) Consejo Superior de Investigaciones Científicas. Revistas científicas electrónicas. E-revistas [Seriada en Internet] junio 2004. Disponible en:<http://www.erevistas.csic.es/especial_revistas/revistas15.htm> Junio 2004.

(30) Oclc.org. WorldCat: Window to the world's library. [Website in Internet]. 2002-10-29. Available from: < http://www.oclc.org/worldcat/default.htm>

(31) Doc6.Consultores en recursos de información - WorldCat y Open WorldCat. Soluciones globales en la Gestión de Información, Documentación y Conocimiento [Sitio Web en Internet]. 2006. Disponible en: < http://www.worldcat.org/&quot>

(32) Oclc.org . What is WorldCat? [Website in Internet].2011. Available from: <http://www.oclc.org/worldcat/policies/privacy>, 2011.

(33) Van de Sompel, H.; Lagoze, C.. "The Santa Fe convention of the Open Archives Initiative". D-Lib[Serials in Internet]. 2000, Feb. v. 6,n.2. Available from: <http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html>

(34) Van de Sompel, H. The Open Archives Initiative Protocol for Metadata Harvesting. [Website in Internet]. 2001-04-24. Available from: <http://www.openarchives.org/OAI/openarchivesprotocol.htm>

(35) Library of Congress. Biblio Tech Review. Z39.50: Part 1- an overview 12 [Website in Internet]. April, 2001. Available from: <http://lcweb.loc.gov/z3950/>

(36) Biblioteca Daniel Cosío Villegas. Bibliotecas en Z39.50 [Sitio Web en Internet]. 2000. Disponible en: <http://biblio.colmex.mx/portal/estela14.JPG>

(37) Cabrera A.M. Coutín A. Las Bibliotecas Digitales. Parte I. Consideraciones teóricas - Características de acceso a la biblioteca digital.. ACIMED [Seriada en Internet] 2005;13 (3). Disponible en: <http://bvs.sld.cu/revistas/aci/vol13_3_05/aci05305.htm>

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

CAPÍTULO 12

12- Obteniendo información mediante la pestaña "descargar" de la interface del bibliotecario de Greenstone

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH)

Palabras clave: Bibliotecas Digitales, Colecciones de documentos., Economía Digital,Estrategias de Inclusión Digital, Greenstone, Sociedad de la Información, Software Libre, UNESCO.

12.1- Introducción

Relacionados con el software Greenstone podemos mencionar varios estándares y lineamientos que son importantes para así poder garantizar un valor presente y futuro de los documentos gestionados y su información, su expansión, integración e interconexión con o migración a otros sistemas. Estos estándares incluyen el Protocolo para Cosecha de Metadata (Protocol for Metadata Harvesting: OAI-PMH), definido por recomendación del Comité Consultivo para Sistemas de Datos Espaciales (CCSDS) y el Dublin Core para definición y estructura de metadatos, entre otros (1).

En este capitulo pretendemos revisar las opciones que muestra Greenstone en la pestaña "descargar" del GLI.

12.2- Bajando información desde Internet vía Web

Esta es la apariencia que muestra la interface del bibliotecario de Greenstone cuando marcamos la pestaña correspondiente a obtener información, la primera a la izquierda destacada. Si hemos decidido obtener información desde la Web esta serán las instrucciones que debemos seguir (2,3). Ver Figura 12.1.

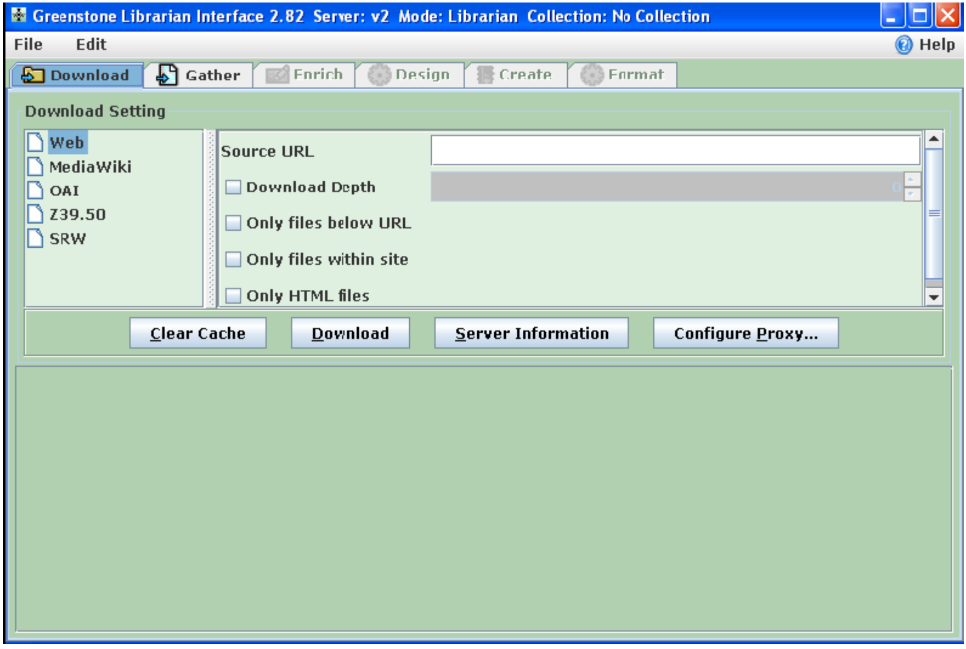


Figura 12.1. Interfaz que caracteriza la pestaña descargar del GLI de Grenstone.

Podemos comenzar utilizando un navegador en Internet y un buscador como Google por ejemplo y situar la dirección: http://englishhistory.net/ tal y como aparece en pantalla en la Figura 12.2.

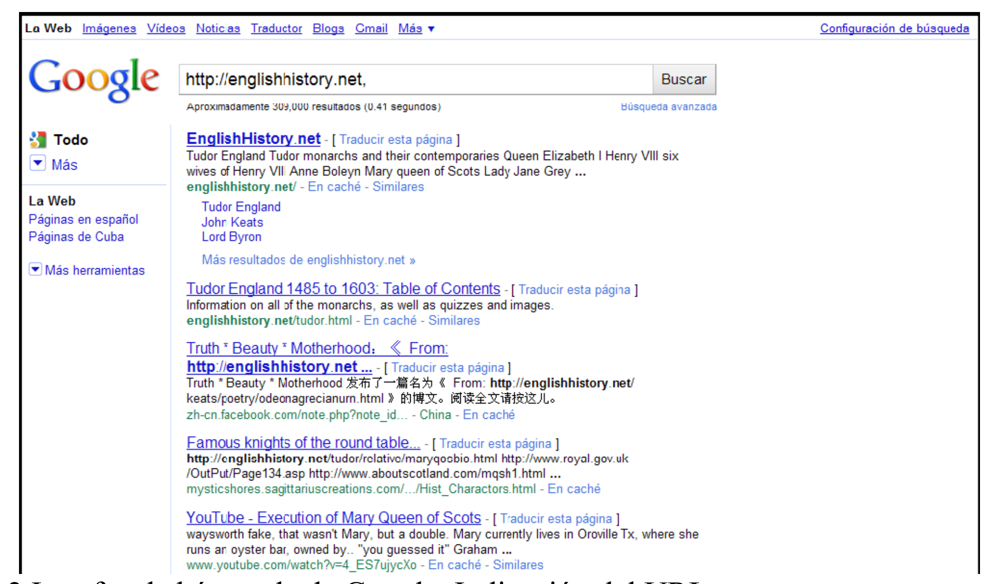


Figura 12.2.Interfaz de búsqueda de Google. Indicación del URL. http://englishhistory.net/tudor.html

Pero mediante la opción Web del Grenstone Ud. puede llegar a este sitio indicado en Google de forma fácil. Ver Figura 12.3.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

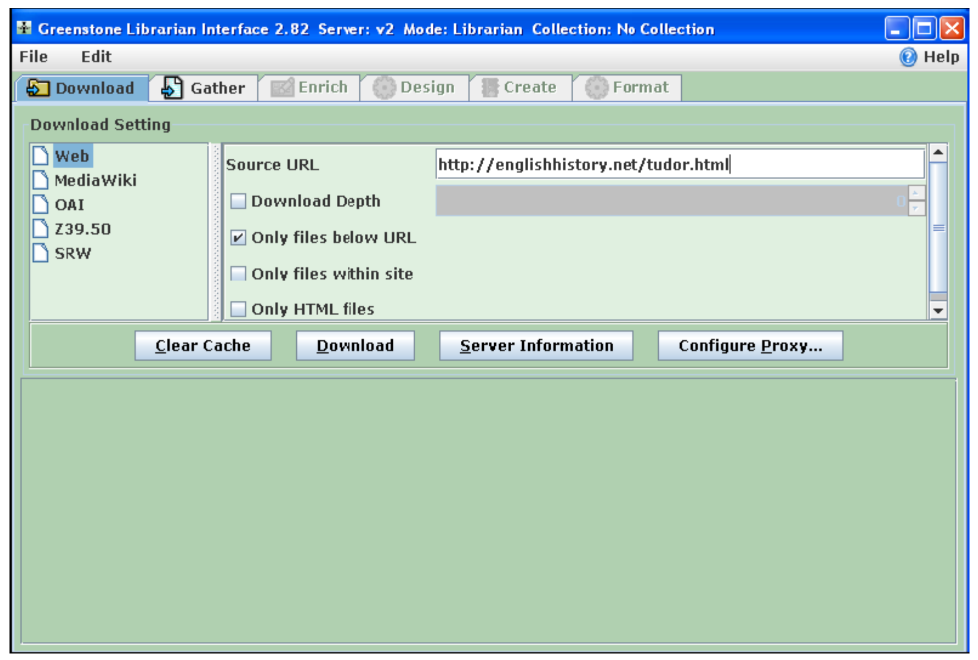


Figura 12.3.Uilización de la opción Web de Greenstone para descargar un sitio Web.

Debemos elaborar algunas notas en la sección de preferencias para lograr resultados. Ver Figura 12.4.

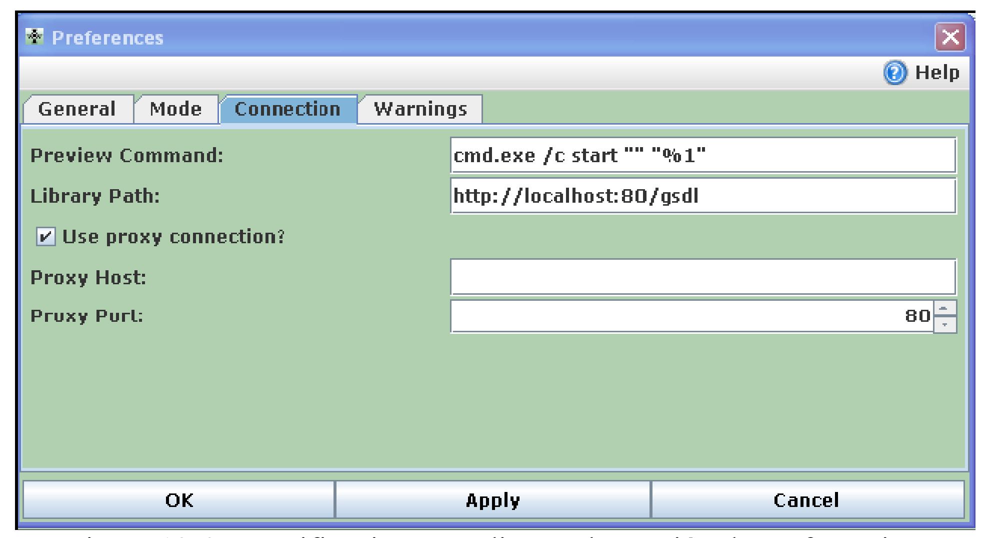


Figura 12.4.Especificaciones a editar en la sección de preferencias.

Si su computadora se encuentra detrás de un firewall o servidor *proxy* Ud. necesitará editar la configuración de su servidor *proxy* en el GLI.

Aparecerá una pequeña ventana donde podrá colocar los datos necesarios. Ahora puede presionar el botón *download* (descargar) y obtener la información requerida y vera una barra que le indicará el progreso de la captura de información.

Los files que haya podido recuperar de la Web se alojarán en una carpeta que aparecerá en el lado izquierdo del panel de trabajo bajo el nombre Downloaded Files , tal y como se observa en la diapositiva en el panel Gather o Reunir en español. Desde allí podrá Ud. Seleccionar entonces los que desee, para incorporarlos en la colección que va a construir. Ver Figura 12.5.

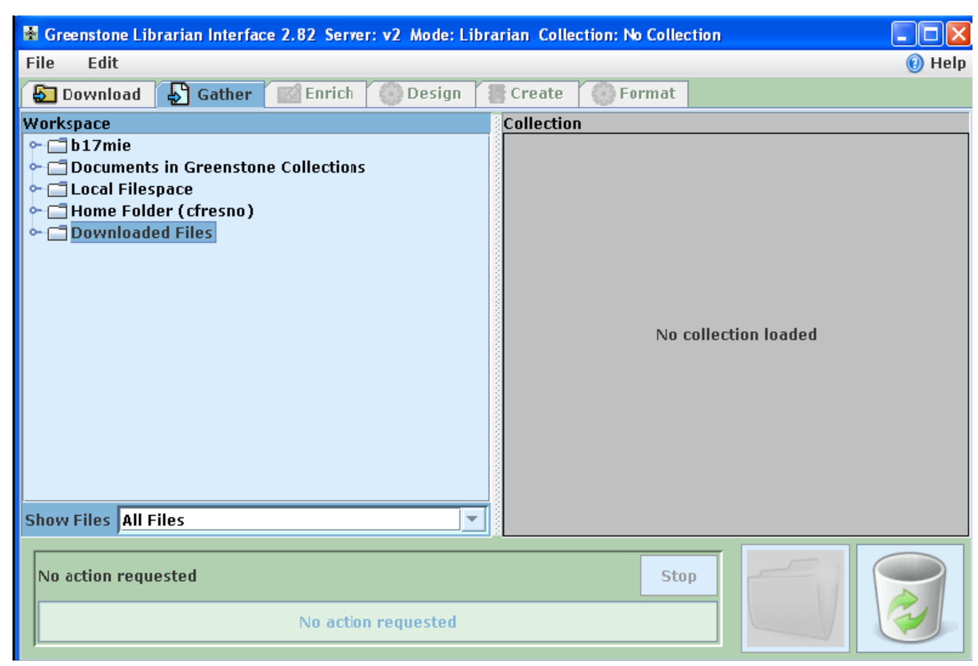


Figura 12.5. Visualización de la carpeta Downloaded Files del espacio de trabajo del GLI procedentes de la Web.

12.3- Características de MediaWiki

MediaWiki es un motor para wikis bajo licencia GNU, programado en PHP. A pesar de haber sido creado y desarrollado para Wikipedia y otros proyectos de la fundación Wikimedia ha tenido una gran expansión desde el año 2005, existiendo un gran número de wikis

basados en este software que nada tienen que ver con dicha fundación, aunque si comparten la idea de la generación de contenidos de manera colaborativa. La mayoría de ellos se dedican a la documentación de software o a temas especializados. Puede ser instalado sobre servidores Web Apache o IIS y puede usar como motor de base de datos MySQL o PostgreSQL (4).Greenstone también le brinda la posibilidad de utilizar esta herramienta y para ello solo debe conocer la dirección de la Wiki de su interés y colocarla en la casilla source URL, como hizo anteriormente con el recurso del Web. Ver Figura 12.6.

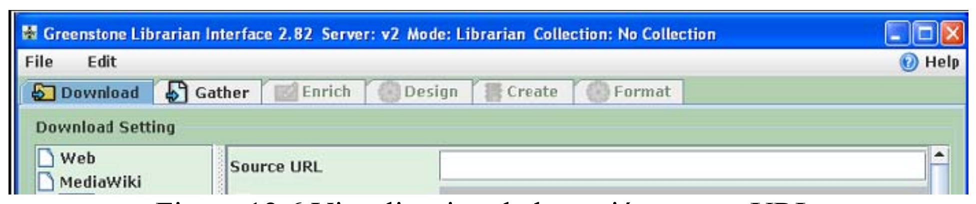


Figura 12.6.Visualizacion de la opción source URL.

12.4- Open Archives Initiative (OAI).Características

Veamos como utilizar el *Open Archive Initiative Protocol for Metadata Harvesting* (OAI-PMH). En esta ocasión no vamos a estar conectados a un proveedor de servicio externo, utilizaremos en su lugar un grupo de files en la forma de XML producidos con el protocolo OAI-PMH (5).

Si quisiéramos conectarnos vía Internet para obtener documentos que utilizan el formato OAI-PMH, pudiéramos introducir en la fuente del URL en forma similar a como lo hicimos previamente con el sitio Web. En la Figura 12.7, hemos utilizado la dirección http://www.doai.org, que conocemos es un servidor OAI.

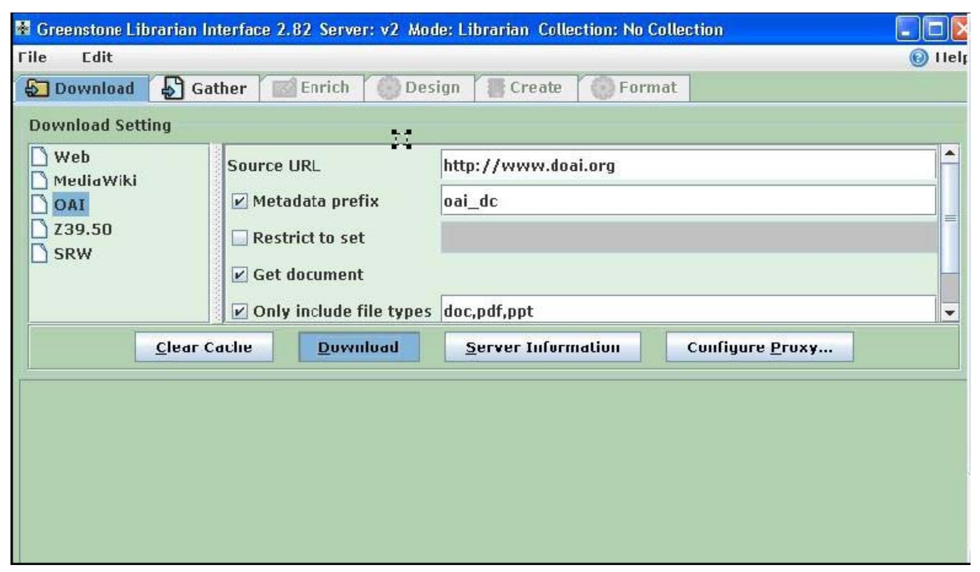


Figura 12.7. Interfaz de descarga de un servidor OAI desde Greenstone.

Marcaríamos la casilla "Get document" que aparece la tercera en el recuadro de la derecha y seguiríamos con los siguientes pasos como en el ejemplo anterior (6). También deberemos especificar que tipos de documentos queremos descargar y lo indicaremos en la casilla "only include file types". Hacemos un clic en "Download" y esperamos los resultados (7). Como hemos trabajado sin conexión a Internet, esto es lo que observaremos, pero queda ejemplificado como el programa hace un seguimiento del nivel de descarga. Ver Figura 12.8.

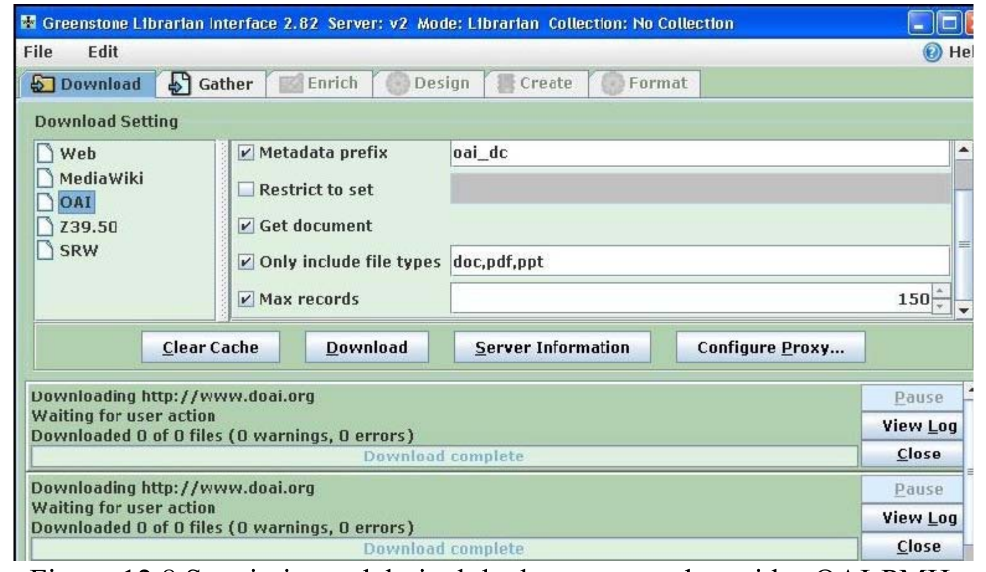


Figura 12.8.Seguimiento del nivel de descarga en el servidor OAI-PMH.

12.5- Open Archives Initiative (OAI) colección y Metadatos

Una vez seleccionados los documentos que vamos a incorporar a la colección, comenzaremos, siguiendo los mismos pasos que para otras colecciones. La denominaremos OAI *Service Provider.* Ver Figura 12.9.

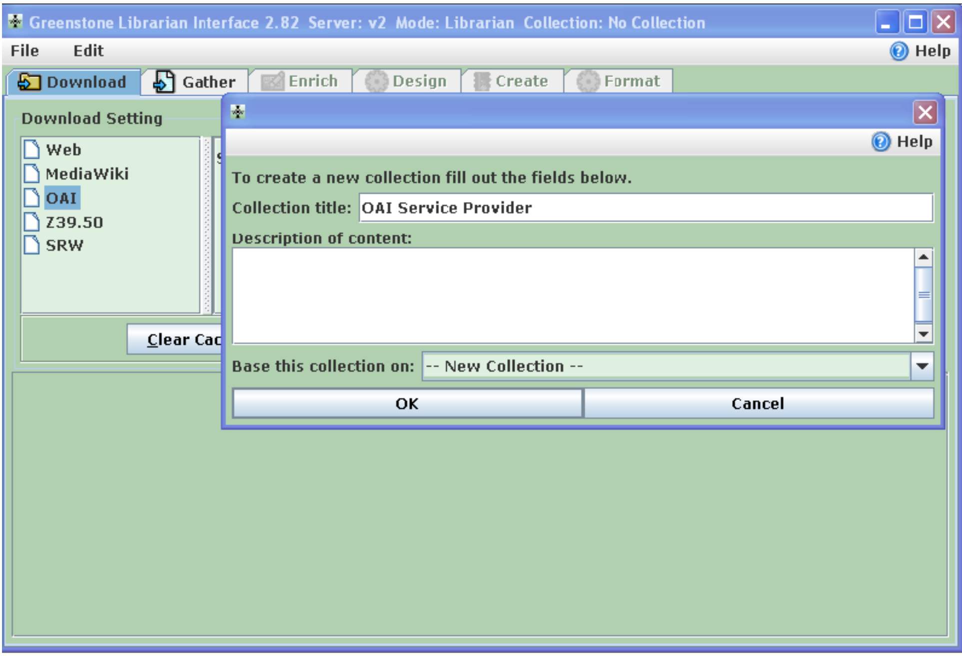


Figura 12.9.Inicio de la creación de la colección de una fuente OAI.

Observaremos en el panel Reunir que tras haber creado la colección esta aparece en el panel izquierdo como se observa en la Figura 12.10.

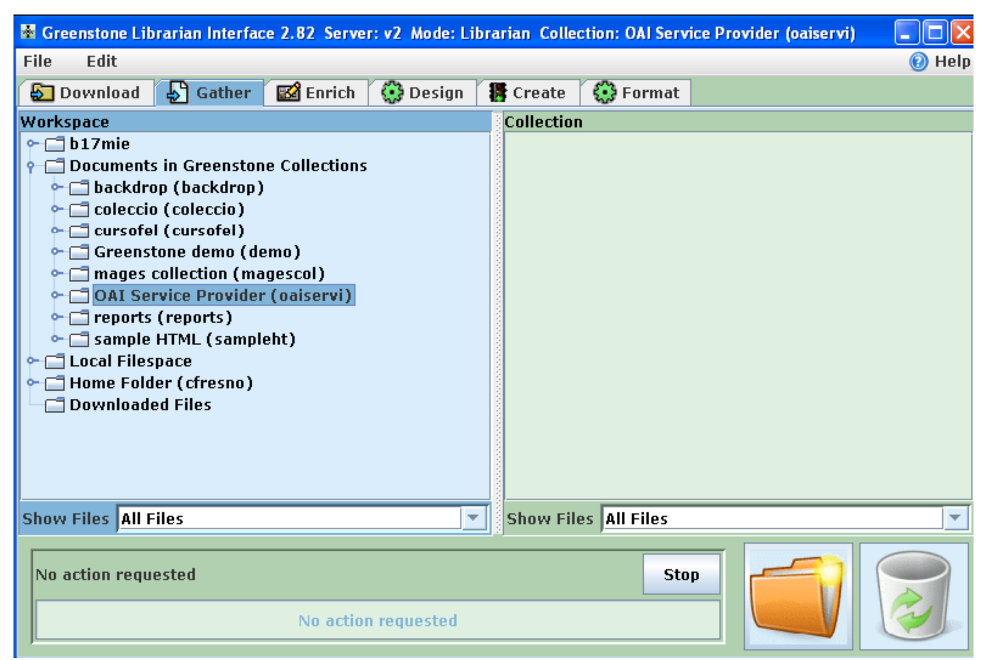


Figura 12.10. Presencia de la colección OAI Service Provider en el panel Reunir.

Es de señalar que mientras Ud. copie la información en esta colección visualizará una ventana que le preguntara si va agregar los plugins OAIPlugin a la lista de los plugins utilizados por la colección porque la interface del bibliotecario no ha encontrado un plugin que pueda manejar esta información. Ver Figura 12.11.

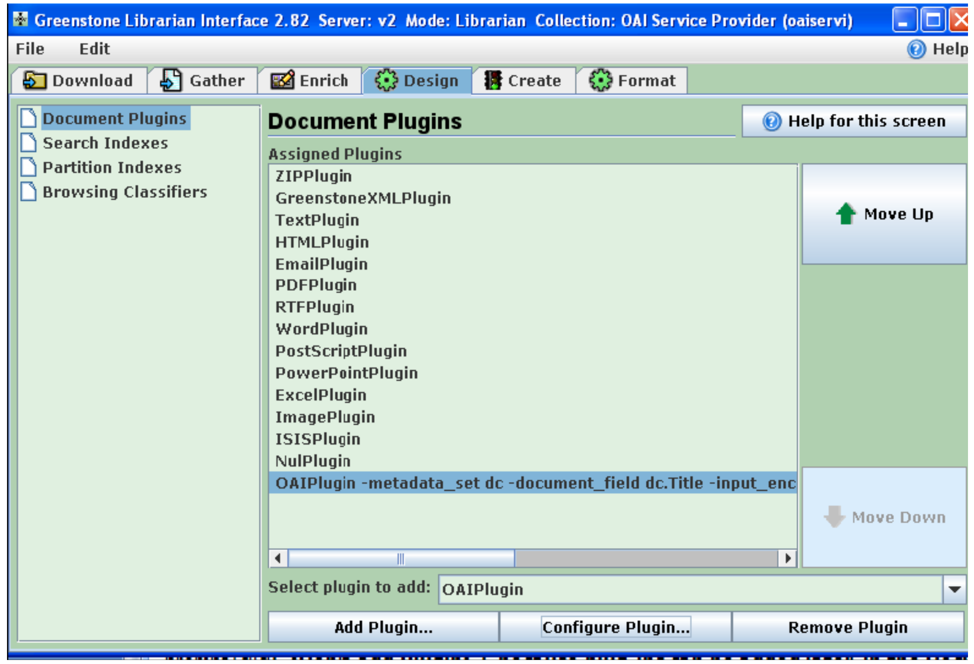


Figura 12.11.Solicitud del programa Greenstone de agregar el OAIPlugin.

Por lo tanto diríjase a la pestaña Diseño localice los plugins necesarios y presione el botón Añadir Plugin. También debe presionar el botón Configurar Plugin. Otro paso necesario con este tipo de colección es seleccionar el ImagePlugin y configurar los argumentos necesarios. El OAIPlugin procesará los documentos OAI y asignará metadatos a las imágenes las cuales serán procesados con ImagePlugin (6). Las imágenes que utiliza Greenstone se almacenan en archivos ubicados en el subdirectorio Greenstone\Web\images (a partir de la versión 2.82), aunque las imágenes de colecciones específicas se pueden conservar en el directorio Greenstone\collect\ nombre de la colección\images (8).

Finalmente una vez creada la colección con este mínimo de requerimientos, esta será la apariencia que tomara en la interfaz del usuario. Ver Figura 12.12

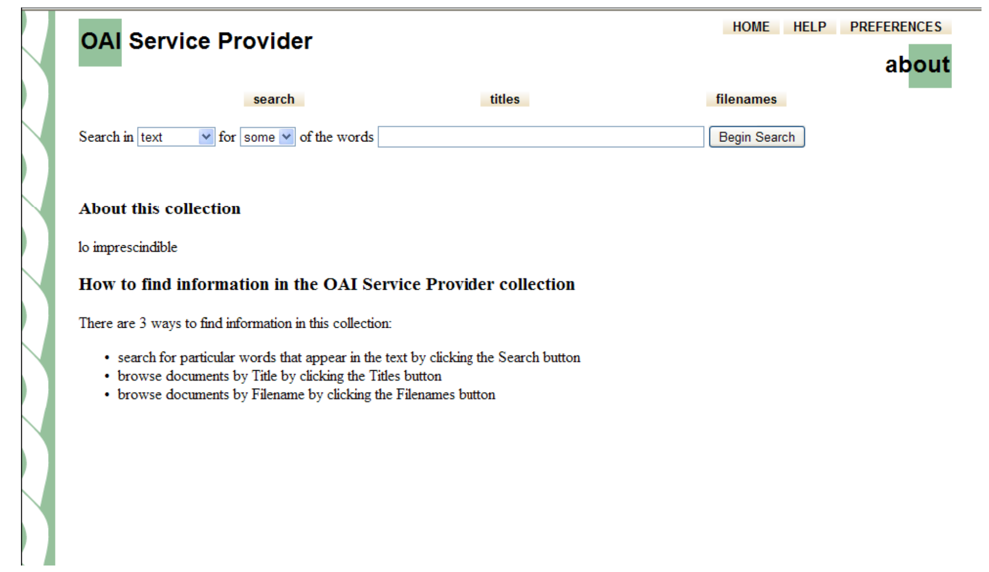


Figura 12.12. Visualización de la colección OAI *Service Provider* en la interfaz de usuario.

12.6- Utilizando la opción Z39.50

Z39.50 es un Estandard internacional que define un protocolo para la recuperación de información entre computadoras. Es necesario que ambas computadoras tengan implementado este protocolo y el usuario puede efectuar la búsqueda sin tener que conocer la sintaxis de búsqueda con que trabaja la computadora en la que se encuentra la información. Z39.50 fue aprobado originalmente por la *National Information Standards Organization (*NISO) en 1988 (9).

Como destacamos en la utilización de la vía Web y OAI para descargar documentos es preciso efectuar algunas notas en la sección de preferencias para lograr un resultado satisfactorio. Ver Figura 12.13.



Figura 12.13.Especificaciones a editar en la sección de preferencias utilizando el standard Z39.50.

En este caso no hemos contado con navegación en Internet, pero la Figura 12.14, ejemplifica como se vería el proceso de descarga.

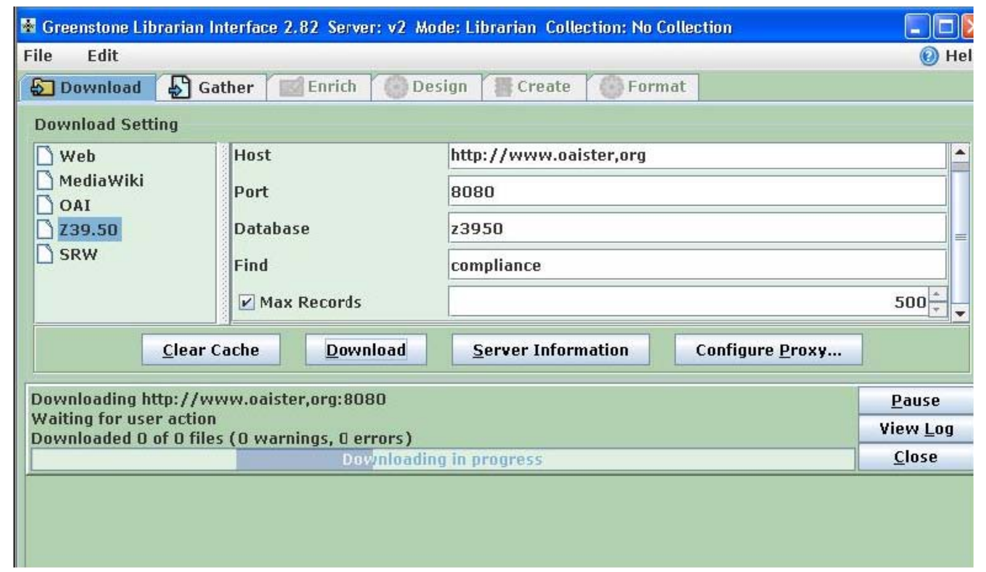


Figura 12.14. Ejemplo del proceso de descarga utilizando el estándar Z39.50.

12.7- Las posibilidades de SRW

SRU (*Search/Retrieve* via URL) es un protocolo de búsqueda normalizado que se utiliza en

Internet para consultar bases de datos mediante el uso de CQL (Commun Query Language), una sintaxis normalizada de consulta. SRW es una variante de SRU en la que los mensajes se dirigen del cliente al servidor no a través de una URL sino utilizando el lenguaje de marcado XML sobre HTTP. SRW, según la recomendación del W3C, utiliza SOAP Simple Object Access Protocol , que especifica cómo encapsular un mensaje dentro de un esquema XML. La norma SRW intenta adherirse a las directrices de la interoperabilidad de servicios Web (10).

Greenstone también cuenta con la posibilidad de trabajar con el protocolo SRW, y procederemos como en los casos anteriores. Ver Figura 12.15.

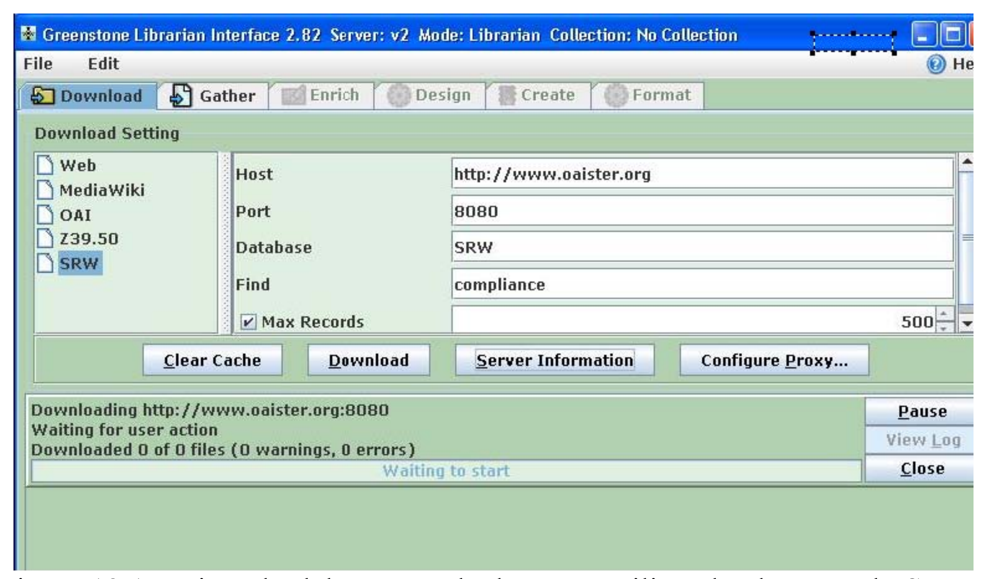


Figura 12.15. Ejemplo del proceso de descarga utilizando el protocolo SRW.

12.8- Las posibilidades de interoperabilidad en las colecciones de Greenstone

Desde un punto de vista informático, interoperabilidad se define como la habilidad que tiene un sistema o producto para trabajar con otros sistemas o productos sin un esfuerzo especial por parte del cliente. Este concepto tiene una importancia creciente a tenor de las colecciones digitales distribuidas que utilizan distintos esquemas de metadatos (11).

Greenstone no solo le permite obtener información desde varios sistemas sino que también resulta posible trasladar las colecciones hechas con Greenstoe a estos otros sistemas. Veamos a continuación como visualizaríamos este fenómeno al seleccionar la opción exportar colección. Ver Figura 12.16.

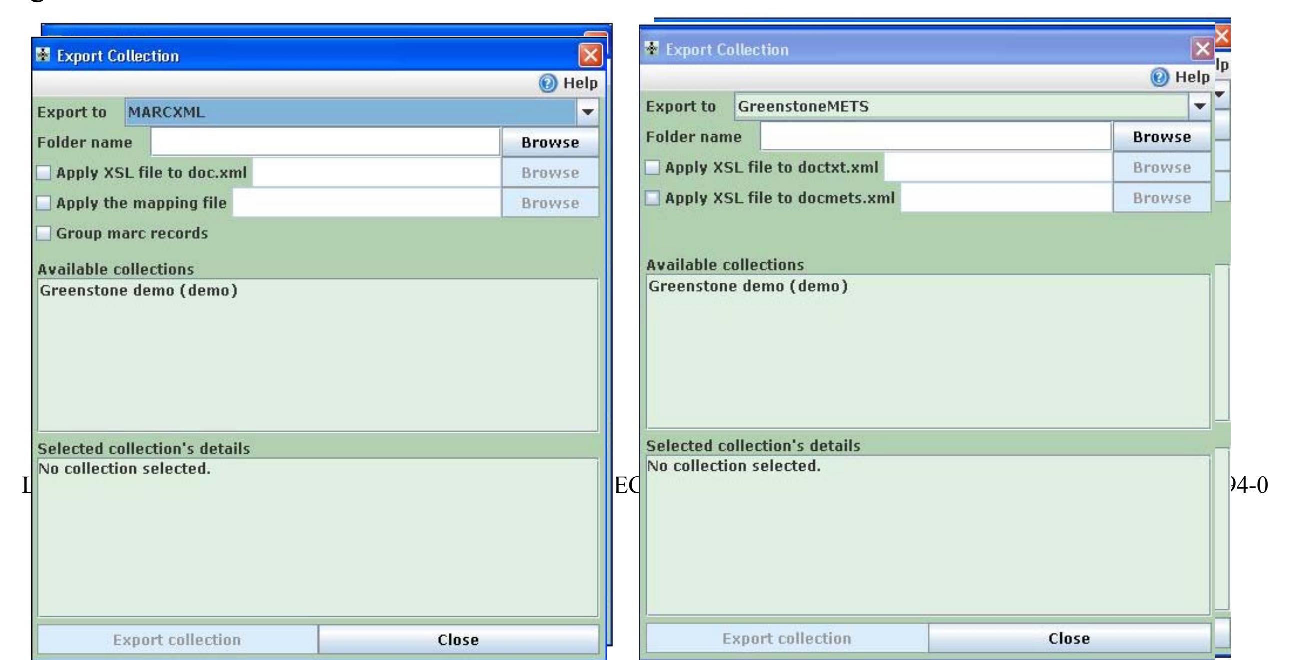


Figura 12.16. Visualización del inicio del proceso de exportación de colecciones en Greenstone hacia DSpace, Fedora, MARCXML, METS.

Una vez concluida la exportación de la colección esta será la pantalla que le indicara que el proceso ha concluido. Ver Figura 12.17.



Figura 12.17. Conclusión del proceso de exportación de la colección DEMO hacia METS.

12.9- Consideraciones Finales.

Existen múltiples funcionalidades explotables con Greenstone y resulta muy notable la incorporación que ha hecho este sistema de una serie de herramientas y aplicaciones que le permiten lograr un nivel de interoperabilidad enorme

12.10- Referencias Bibliográficas

(1) Miranda C. Sistemas de Gestión de Documentos bajo Licencia de Código Libre [Sitio Web en Internet] 08 Junio 2009. Disponible en:<http://opengraphprotocol.org/schema/xmlns:fb>

(2) Greenstone .org. Lab 10: Miscellaneous [Website in Internet] 2009. Available from: <http://www.greenstone.org/Lab10>

(3) Bainbridge D, McKay D, Witten H. I. Guía del programador. Biblioteca digital Greenstone. [Sitio Web en Internet].Departamento de Informática. Universidad de Waikato, Nueva Zelanda. Marzo de 2004.Disponible en:<http://www.greenstone.org/manuals/gsdl2/es/html/Develop_es_all.html >

(4) Wikipedia. MediaWiki [Website in Internet] 3 June 2010. Available from: <http://es.wikipedia.org/apple-touch-icon.png >

(5) Openarchives.org. The Open Archives Initiative Protocol for Metadata Harvesting. Protocol Version 1.0 of 2001-01-21. [Website in Internet] 2001-04-24. Available from: <http://www.openarchives.org/OAI/openarchivesprotocol.htm >

(6) Greenstone .org. Lab 9: Sharing collections with OAI-PMH [Website in Internet] 2009. Available from: <http://www.greenstone.org/Lab9 >

(7) Wiki.greenstone.org .Greenstone tutorial exercise. Downloading over OAI 2010[Website in Internet] 2010. Available from: <http://wiki.greenstone.org/wiki/gsdoc/tutorial/en/OAI_collection.htm >

(8) Grenstone.org. Biblioteca Digital Greenstone. Guía del Programador. Capitulo 3. El Sistema de Ejecución de Greenstone [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css >

(9) Library of Congress. About Z39.50 Gateway. [Website in Internet] August 29, 2011. Available from: < http://www.loc.gov/z3950/gateway.html >

(10) Library of Congress. Search Retrieval/via URL. Official Web Site [Website in Internet].2000. Available from: <http://www.loc.gov/standards/sru >

(11) Méndez E. Senso A. Introducción a los metadatos. Estándares y aplicación Conceptos asociados al uso de meta información . Unidad de Autoformación.[Sitio Web en Internet] SEDIC 2004. Disponible en: <http:www.sedic.es >

CAPÍTULO 13

13- Configuración del servidor de Greenstone para usuarios

Ms C Ricardo Fernández Rodríguez (ricardo@cecam.sld.cu) Profesor Auxiliar del I.S.C.M.H. Máster en Informática en Salud. Centro de Cibernética Aplicada a la Medicina (CECAM) I.S.C.M.H. Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba. CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH)

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH)

Palabras clave: Bibliotecas Digitales, Desarrollo de software. Desarrollo Sostenible, Economía Digital, Estrategias de Inclusión Software Libre, UNESCO.

13.1- Introducción

Las versiones mas recientes de Greenstone (a partir de la 8.2) cuentan con un servidor Web incorporado en el paquete de instalación. Pero necesitamos además un servidor Web externo, si pretendemos enlazar varias computadoras.

El servidor Web Apache de http://www.apache.org.es libre y fácil de instalar (1) se trata de un servidor *open source* para los modernos sistemas operativos que incluye UNIX y Windows NT. El objetivo es proveer un servidor http, seguro, eficiente, y extensible con los estándares http vigentes (2).

En el caso de instalar sobre Linux, la instalación deberá hacerse desde una cuenta de administrador o root; para el caso de Windows debe hacerse desde el administrador o un usuario del grupo de administradores. Aunque puede instalarse sin ser administrador, es conveniente hacerlo correctamente desde un principio para evitar inconvenientes posteriores con los permisos.

13.2- Flexibilidad de uso de Greenstone

Greenstone corre en todas las versiones de Windows, Unix/Linux, y Mac OS-X. Es de muy fácil instalación. Para la instalación en Windows no es necesaria ninguna configuración especial, el usuario final puede instalar el software en su computadora personal. Los usuarios institucionales pueden instalar el software en un servidor Web y entonces interoperar con los estándares Web del software servidor (3).Recordemos de capítulos anteriores que en informática (4), un servidor es un tipo de software que realiza ciertas tareas en nombre de los usuarios. El término servidor ahora también se utiliza para referirse al ordenador físico en el cual funciona ese software, una máquina cuyo propósito es proveer datos de modo que otras máquinas puedan utilizar esos datos (para mayor ampliación sobre el término consulte el Capitulo 18 Términos Conceptuales y Referenciales).

13.3- Biblioteca local o biblioteca en el Web de Greenstone

La mayor diferencia entre las dos es que resulta posible instalar una biblioteca local que integre unas pocas computadoras pues el Greenstone posee su propia aplicación Web server. Mientras que la biblioteca en el Web requiere de un servidor Web externo, que puede ser el Apache o el Microsoft IIS, pero se recomienda el primero (5).

Y se recomienda además que Ud. Instale la biblioteca local antes de proceder a instalar el servidor Web (5).Los pasos necesarios para instalar la biblioteca local han sido revisados en el capitulo 5 de este libro, les sugerimos revisar esto antes de proceder con la instalación del Web server.

13.3.1- ¿Cómo se hará accesible mi servidor Web a Internet?

Para esto, necesitará de una computadora que esté permanentemente conectada a Internet y el Web server corriendo (5). Si no cuenta con este tipo de computadora será necesario que obtenga un servicio de hosting apropiado que le garantice que los ejecutables de CGI puedan correr apropiadamente

13.4- Operaciones básicas de configuración

Para configurar el servidor WEB primeramente tenemos que conocer en que disco y carpetas esta instalado el Greenstone (recuerde que previamente hemos instalado el modo local) y donde instalamos el servidor apache.

Por ejemplo supongamos que el Greenstone lo tenemos instalados en la torre D y en una carpeta llamada Biblioteca (d:/ Biblioteca) y el apache en la torre D y en una carpeta (d:/Appeserv).

Nos dirigimos a la carpeta donde se ubica el Apache y entramos en la carpeta apache2.2 y dentro de esta carpeta entramos en la carpeta conf y este seria el camino (Ver Figura 13.1) d:/appeserv/apache2.2/conf

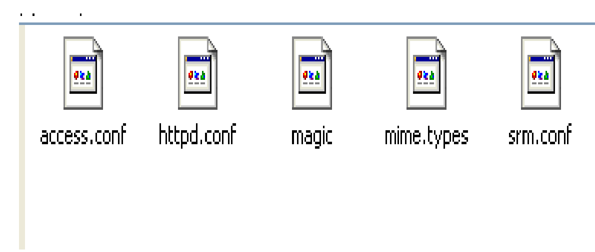
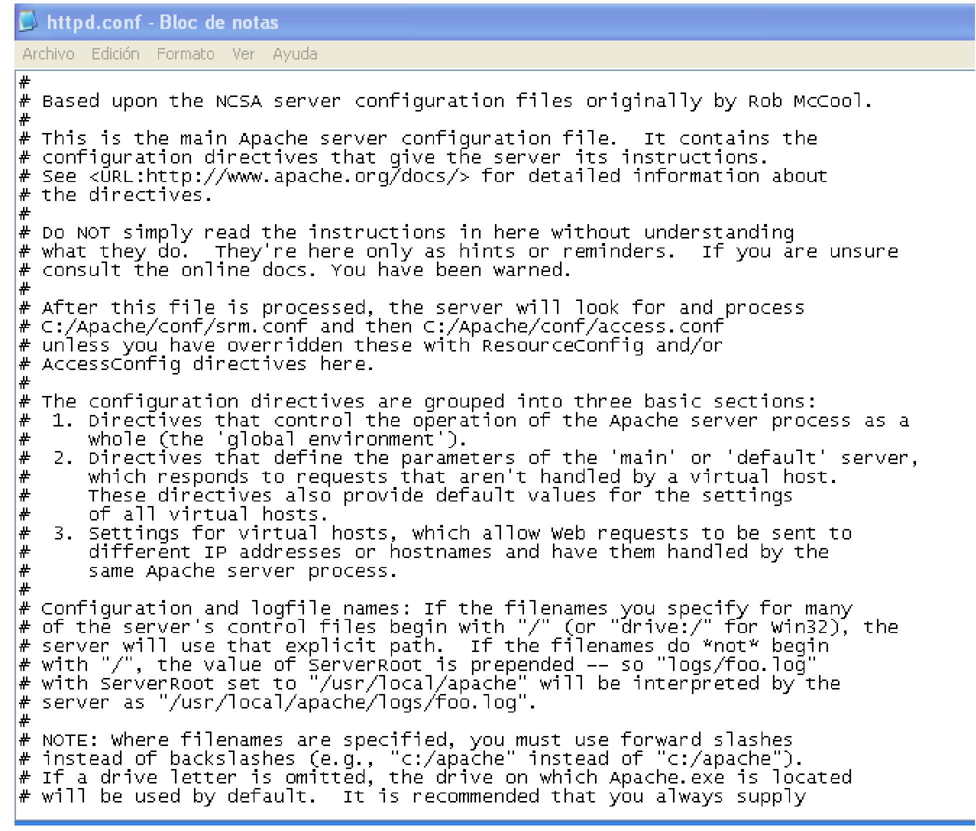


Figura 13.1-Contenido de la carpeta conf de Apache.

Ya en la carpeta conf abrimos utilizando el notepad el archivo httpd.conf (Ver Figura 13.2) y al final del mismo copiamos las siguientes instrucciones (6):

```
ScriptAlias /Greenstone/cgi-bin "c:/Greenstone/cgi-bin"
 <Directory "/Greenstone/cgi-bin">
 Options None
 AllowOverride None
 </Directory>
Alias /Greenstone "c:/Greenstone"
 <Directory "c:/Greenstone">
 Options Indexes MultiViews FollowSymLinks
 AllowOverride None
 Order allow,deny
 Allow from all
 </Directory>
```



```
httpd.conf - Bloc de notas
Archivo  Edición  Formato  Ver  Ayuda
#
# Based upon the NCSA server configuration files originally by Rob McCool.
#
# This is the main Apache server configuration file.  It contains the
# configuration directives that give the server its instructions.
# See <URL:http://www.apache.org/docs/> for detailed information about
# the directives.
#
# Do NOT simply read the instructions in here without understanding
# what they do.  They're here only as hints or reminders.  If you are unsure
# consult the online docs. You have been warned.
#
# After this file is processed, the server will look for and process
# C:/Apache/conf/srm.conf and then C:/Apache/conf/access.conf
# unless you have overridden these with ResourceConfig and/or
# AccessConfig directives here.
#
# The configuration directives are grouped into three basic sections:
#  1. Directives that control the operation of the Apache server process as a
#     whole (the 'global environment').
#  2. Directives that define the parameters of the 'main' or 'default' server,
#     which responds to requests that aren't handled by a virtual host.
#     These directives also provide default values for the settings
#     of all virtual hosts.
#  3. Settings for virtual hosts, which allow Web requests to be sent to
#     different IP addresses or hostnames and have them handled by the
#     same Apache server process.
#
# Configuration and logfile names: If the filenames you specify for many
# of the server's control files begin with "/" (or "drive:/" for Win32), the
# server will use that explicit path.  If the filenames do *not* begin
# with "/", the value of ServerRoot is prepended -- so "logs/foo.log"
# with ServerRoot set to "/usr/local/apache" will be interpreted by the
# server as "/usr/local/apache/logs/foo.log".
#
# NOTE: Where filenames are specified, you must use forward slashes
# instead of backslashes (e.g., "c:/apache" instead of "c:\apache").
# If a drive letter is omitted, the drive on which Apache.exe is located
# will be used by default.  It is recommended that you always supply
```

Figura 13.2. Contenido del archivo httpd.conf abierto con el programa notepad

Es recomendable editar estas instrucciones, sobre todo si tenemos instalado el Greenstone en el disco D en una carpeta llamada Biblioteca y en estas instrucciones observamos que se hace referencia a una torre de disco C y a una carpeta llamada Greenstone. En este caso procederíamos del modo siguiente; tendríamos que cambiar la torre de disco C por D y la carpeta Greenstone por la carpeta Biblioteca, quedando estas instrucciones así (6):

```
ScriptAlias /Biblioteca/cgi-bin "d:/Biblioteca /cgi-bin"
<Directory "/Biblioteca/cgi-bin">
Options None
AllowOverride None
</Directory>

Alias /Greenstone "d:/Biblioteca"
<Directory "d:/Biblioteca">
Options Indexes MultiViews FollowSymLinks
AllowOverride None
Order allow,deny
Allow from all
</Directory>
```

Los nombres de las carpetas deben de mantener sus minúsculas y mayúsculas tal como están en el directorio ya que apache es sensitivo a cambios de mayúsculas por minúsculas y viceversa, por ejemplo si la carpeta se llama Biblioteca no puedo poner en la configuración "biblioteca".

Es también importante conocer y comentaremos ahora, que si se trata del sistema operativo Linux todos los nombres de carpetas y archivos tienen que estar en minúscula.

Una vez editado este archivo httpd.conf lo guardamos y entonces pasamos a reinicializar el servidor

para que todos estos cambios sufran efecto.

Si necesitamos detener el servidor entramos en la carpeta AppServ y posteriormente entramos en la carpeta Apache2.2 y seguidamente en la carpeta bin (Ver Figura 13.3).



Figura 13.3. Contenido de la carpeta bin de Apache

Dentro de esa carpeta ejecutamos el archivo "ApacheMonitor.exe" (ver primer ícono de la Figura 13.13); una vez ejecutado este archivo cambia el color, este programa se inicializa y lo constatmos en la barra de tareas y podemos de este modo detener o inicializar el servidor, por ejemplo al hacer un clic en él podemos detener el servidor (Ver Figura 13.4).

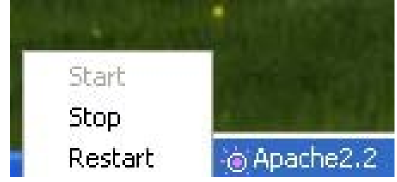


Figura 13.4. Inicialización en la barra de tareas del Apache monitor.

Al hacer clic en la opción stop vemos como el icono toma otra apariencia (se pone en rojo el ícono que identifica al "ApacheMonitor" lo cual significa que el servidor está detenido). Al teclear en el navegador http://localhost observaríamos la siguiente pantalla en la cual podemos verificar que el servidor no esta funcionando (Ver Figura 13.5).

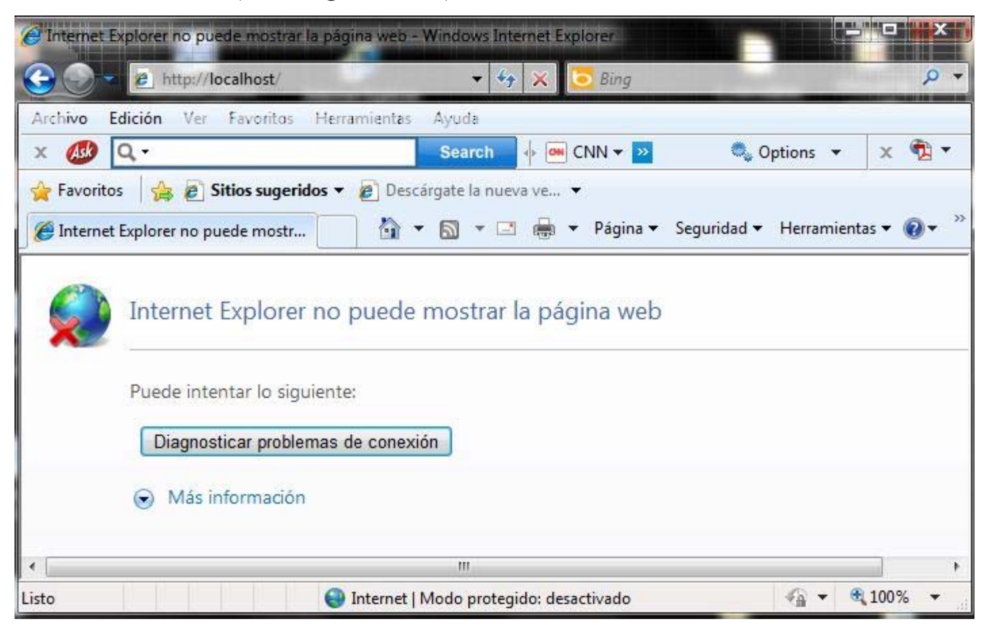


Figura 13.5. Comprobación del no funcionamiento del servidor Apache.

Posteriormente reinicializamos el servidor para que los cambios realizados en el archivo http.conf se hagan efectivos para lo cual de nuevo vamos a la barra de inicio y ejecutamos el ícono "ApacheMonitor" y al dar clic en la opción start entonces se inicializa el servidor (Ver Figura 13.6). y el icono tomaría esta apariencia (constatamos que se pone en verde lo que significa que el

servidor esta inicializado). y al teclear en el navegador *http://localhost* obtendríamos la siguiente pantalla la cual nos muestra como el servidor esta funcionado (Ver Figura 13.7).

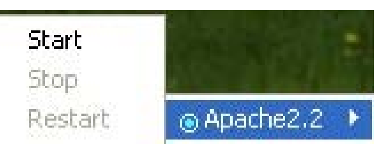


Figura 13.6. Opción start del monitor Apache.

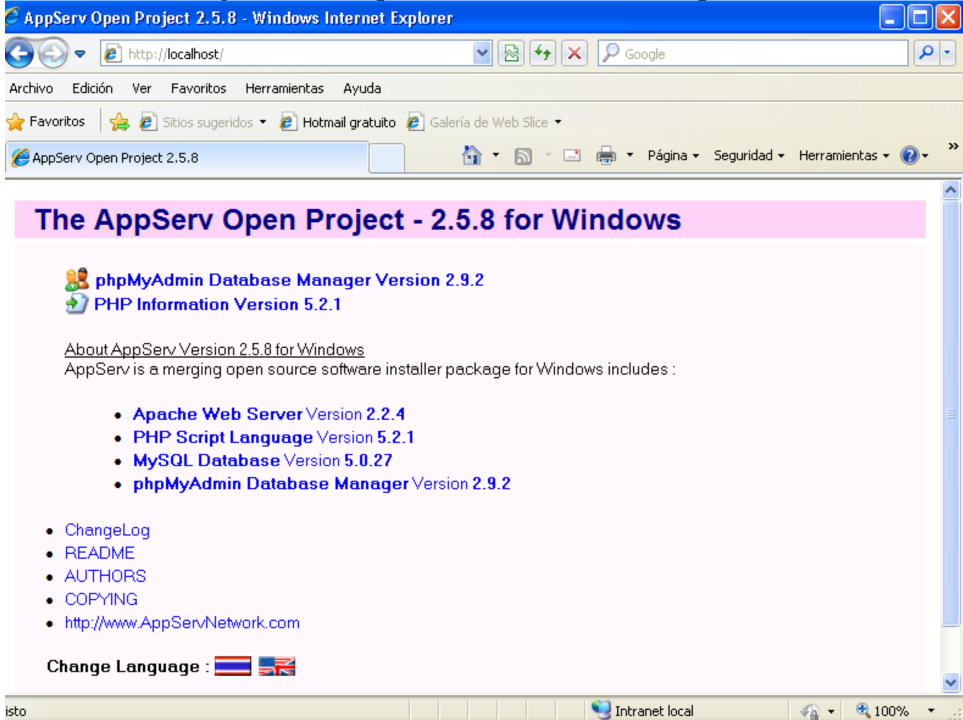


Figura 13.7. Interfaz inicial de trabajo del software Apache.

Posteriormente veamos si es posible conectarse vía Web con la carpeta de Greenstone instalada que llama Biblioteca y para esto teclearemos http://localhost/Biblioteca y se nos presenta la siguiente pantalla (Ver Figura 13.8):

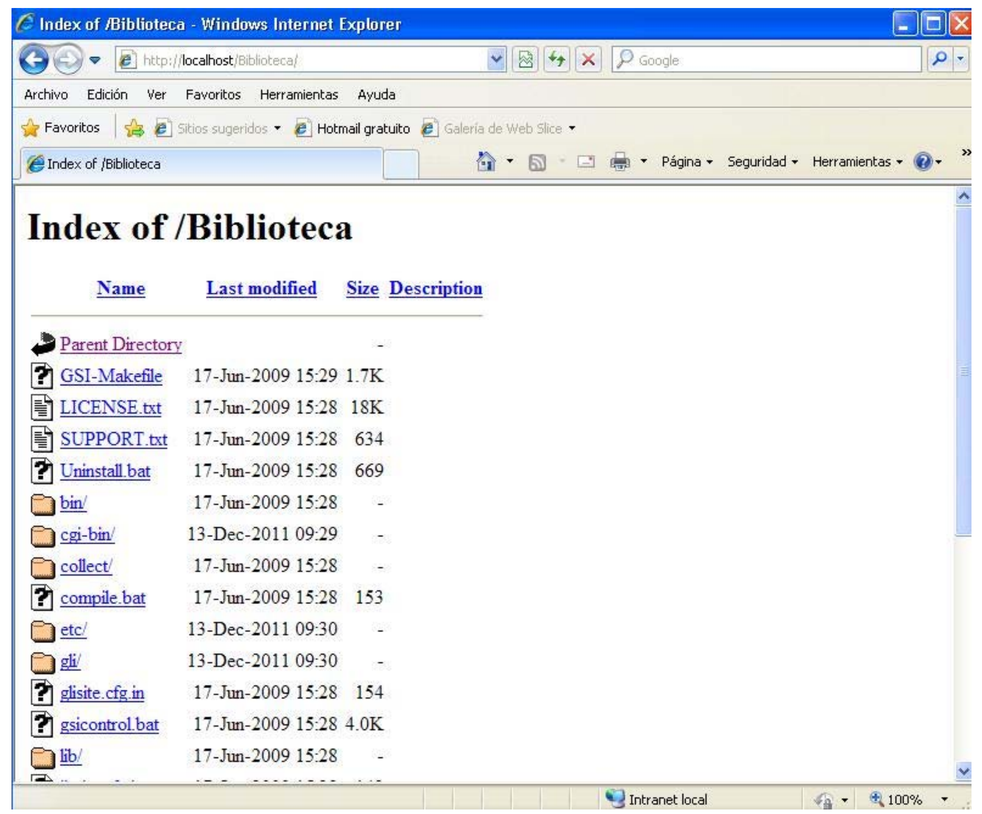


Figura 13.8. Pantalla localhost en el Web del programa Greenstone.

En esta pantalla vemos todos los archivos y carpetas que conforman el Greenstone que se encuentra en la carpeta Biblioteca sin embargo aun no hemos entrado a la biblioteca. Si tecleamos la dirección http://localhost/Biblioteca/cgi-bin/library.cgi se nos presentará la siguiente pantalla (Ver Figura 13.9):

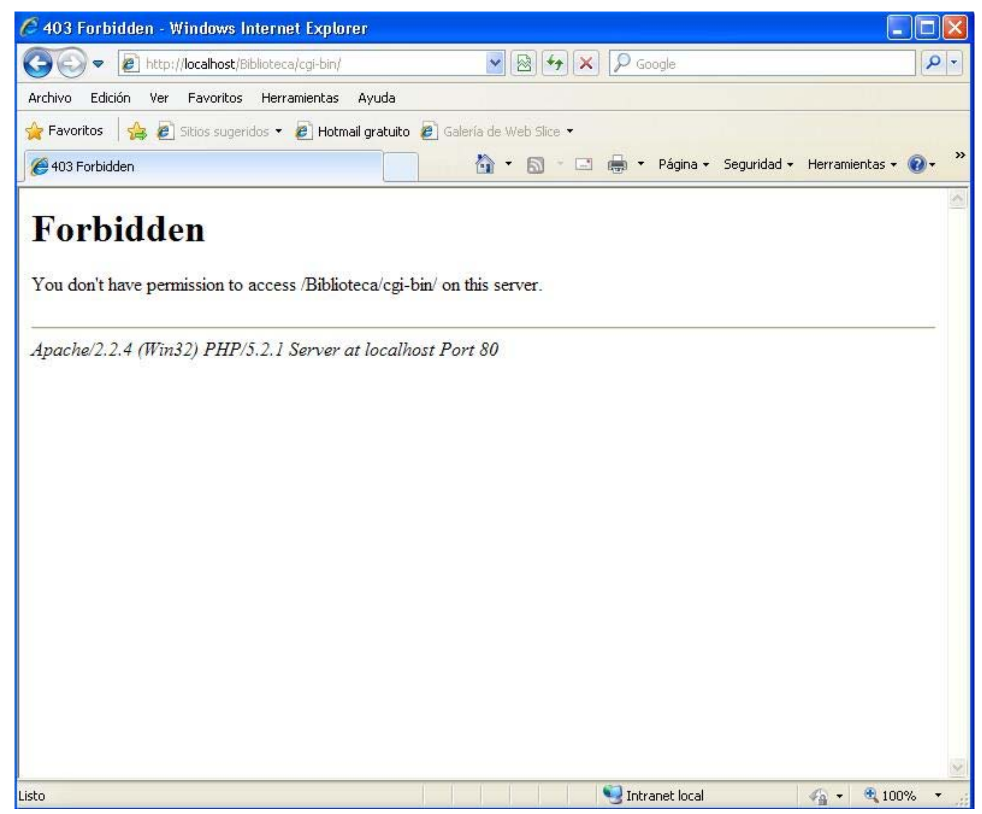


Figura 13.9. Pantalla que indica el fracaso a ingresar en la biblioteca Web.

Esto se debe a que aun no hemos configurado el archivo gsdlsite.cfg.in

Lo primero que debemos hacer es quitarle la extensión .in y dejarlo como gsdlsite.cfg y pasar a editar 4 líneas del contenido del mismo como se muestra a continuación (6):

```
# this file should be placed in the same directory as your library
# executable file. it defines parameters that are particular to a
# given site, and therefore should be edited to suit your site.

# points to the GSDLHOME directory
# This must be set, and must point to the top level Greenstone directory
# (usually gsdl or Greenstone)
gsdlhome **GSDLHOME**

# this is the http address of GSDLHOME
# if your webserver's DocumentRoot is set to $GSDLHOME
# then httpprefix can be commented out
httpprefix /greenstone

# this is the http address of the Greenstone Web directory (which contains
# images, style etc)
# if your Web directory is a top level folder inside $GSDLHOME, then this
# can remain commented out, and will default to httpprefix/Web
#httpweb /greenstone/Web

# this should contain the http address of the library.cgi script. This
```

```
# is not needed if the http server sets the environment variable
# SCRIPT_NAME
# Generally this is not needed with Apache
#gwcgi /greenstone/cgi-bin/library.cgi

# maxrequests is the most requests a fastcgi process
# will serve before it exits. This can be set to a
# low figure (like 1) while debugging and then set
# to a high figure (like 10000) when everything is
# working well.
# Only applies if you are using fastcgi
maxrequests 10000

# Remote Greenstone Options (for server side)

# If Perl is not already on your PATH, set the following to the
# full path of the perl bin directory (Perl's parent folder)
# and it will be prepended to the PATH.
#perlpath /usr/bin

# Options for FLI

# If using Greenstone's GLIServer to run the Fedora Librarian
# Interface (FLI), then you need to set fedora version and fedora home
# variables, either below, or as environment variables.
# FEDORA_VERSION is the major version number of Fedora and
# FEDORA_HOME is the full path to your Fedora installation.
# Setting fedorahome and fedoraversion here will over-ride the
# FEDORA_HOME and FEDORA_VERSION environment variables.
#fedorahome /full/path/to/fedora
#fedoraversion 3

# If using Greenstone's GLIServer to run the Fedora Librarian
# Interface (FLI), then you need to set the full path to Java either as below
# or as the JAVA_HOME environment variable.
# Setting javahome here will over-ride the JAVA_HOME environment variable.
#javahome /full/path/to/j2sdk1.4-or-higher
```

Las líneas que tenemos que editar están en negrita, itálica y a un tamaño mayor esta son:

gsdlhome **GSDLHOME**
httpprefix /greenstone
#httpweb /greenstone/Web
#gwcgi /greenstone/cgi-bin/library.cgi

Por lo tanto tenemos que poner los caminos de acuerdo a donde instalamos el Greenstone, en este caso lo instalamos en una carpeta llamada Biblioteca por los tanto quedarían de la siguiente forma la primera.

gsdlhome "/Biblioteca"

la segunda:

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

httpprefix /Biblioteca

tercera:

httpweb /Biblioteca/Web

la cuarta:

gwcgi /Bibliotrca/cgi-bin/library.cgi

por tanto el contenido del archivo *gsdlsite.cfg* quedaría de la siguiente forma (7):

```
# this file should be placed in the same directory as your library
# executable file. it defines parameters that are particular to a
# given site, and therefore should be edited to suit your site.

# points to the GSDLHOME directory
# This must be set, and must point to the top level Greenstone directory
# (usually gsdl or Greenstone)
gsdlhome "/Biblioteca"

# this is the http address of GSDLHOME
# if your webserver's DocumentRoot is set to $GSDLHOME
# then httpprefix can be commented out
httpprefix /Biblioteca

# this is the http address of the Greenstone Web directory (which contains
# images, style etc)
# if your Web directory is a top level folder inside $GSDLHOME, then this
# can remain commented out, and will default to httpprefix/Web
httpweb /Biblioteca/Web

# this should contain the http address of the library.cgi script. This
# is not needed if the http server sets the environment variable
# SCRIPT_NAME
# Generally this is not needed with Apache
gwcgi /Bibliotrca/cgi-bin/library.cgi

# maxrequests is the most requests a fastcgi process
# will serve before it exits. This can be set to a
# low figure (like 1) while debugging and then set
# to a high figure (like 10000) when everything is
# working well.
# Only applies if you are using fastcgi
maxrequests 10000

# Remote Greenstone Options (for server side)

# If Perl is not already on your PATH, set the following to the
# full path of the perl bin directory (Perl's parent folder)
# and it will be prepended to the PATH.
#perlpath /usr/bin

# Options for FLI

# If using Greenstone's GLIServer to run the Fedora Librarian
```

```
# Interface (FLI), then you need to set fedora version and fedora home
# variables, either below, or as environment variables.
# FEDORA_VERSION is the major version number of Fedora and
# FEDORA_HOME is the full path to your Fedora installation.
# Setting fedorahome and fedoraversion here will over-ride the
# FEDORA_HOME and FEDORA_VERSION environment variables.
#fedorahome /full/path/to/fedora
#fedoraversion 3

# If using Greenstone's GLIServer to run the Fedora Librarian
# Interface (FLI), then you need to set the full path to Java either as below
# or as the JAVA_HOME environment variable.
# Setting javahome here will over-ride the JAVA_HOME environment variable.
#javahome /full/path/to/j2sdk1.4-or-higher
```

Ahora bien si tecleamos en el navegador http://localhost/Biblioteca/cgi-bin/library.cgi se nos presentaría el sitio Web de la biblioteca digital Greenstone (Ver Figura 13.10).

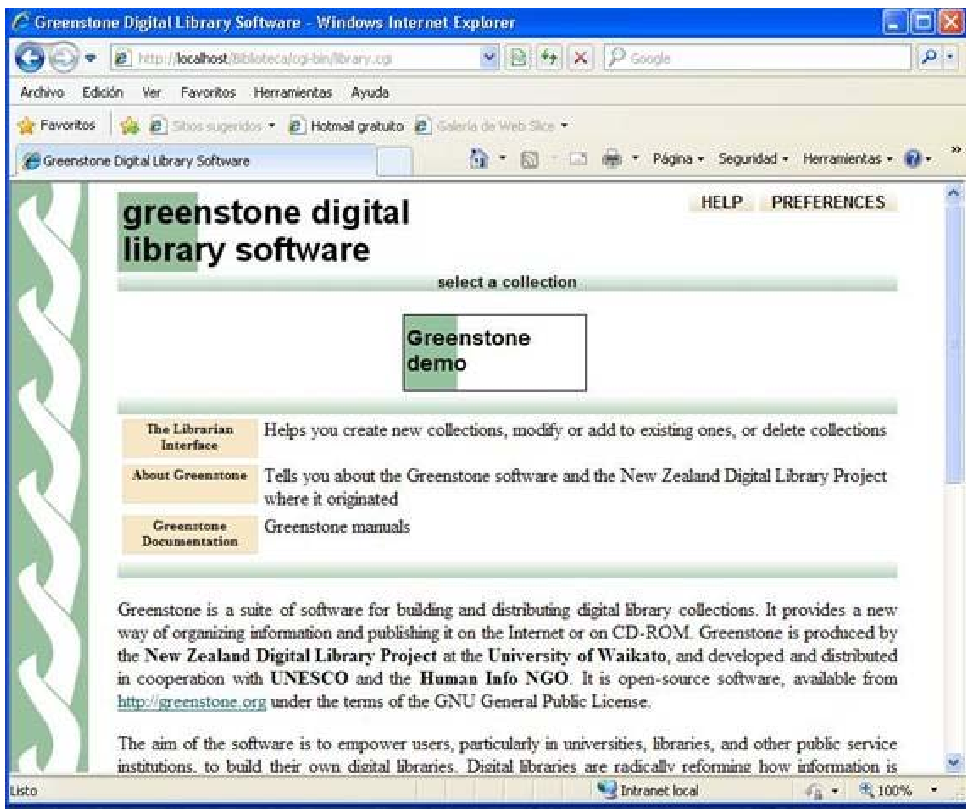


Figura 13.10. Interfaz inicial de la biblioteca Greenstone tras la modificación del archivo gsdlsite.cfg.in.

Ya con todos estos pasos anteriores queda configurado el sitio Web de la biblioteca digital Geenstone el cual puede ser accedido desde cualquier maquina de una intranet siempre y cuando sustituyamos en la dirección localhost por el nombre de la maquina o por el número IP de la maquina.

Por ejemplo supongamos que la maquina donde se instalo la biblioteca se llama server y su número IP es 192.168.0.46 entonces desde cualquier maquina de la intranet tecleando http://server/Biblioteca/cgi-bin/library.cgi o http://192.168.0.46 /Biblioteca/cgi-bin/library.cgi se nos presenta el sitio Web de la biblioteca digital Geenstone.

13.5- Consideraciones Finales

Como hemos podido demostrar no presenta una gran complejidad la configuración del servidor Web para colocar nuestra biblioteca en Internet. Solo es precisa que sigan al pie de la letra las instrucciones brindadas y les aseguramos que tendrán éxito en la tarea.

13.6- Referencias Bibliográficas

(1) Greenstone.org. Guía de Instalación Greenstone. Apendice: Programas Asociados.[Sitio Web en Internet]. 2007. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone>

(2) Apache. Http Server Project. The number one HTTP server on the Internet.[Website in Internet]2010/04/27. Available from: <http://httpd.apache.org>

(3) Greenstone.org Factsheet :Greenstone Digital Library Software [Website in Internet] 11/2000. Available from: <http://www.greenstone.org/factsheet>

(4) Pergamino Virtual.com. Definición de Servidor [Sitio Web en Internet]. 11/08/10. Disponible en: <http://www.pergaminovirtual.com/definicion/default.css>

(5) Wiki.greenstone.org. Installing Greenstone.[Website in Internet] 17 August 2009.Available from: <http://wiki.greenstone.org/> greenstone wiki <17 August 2009>

(6)-Barrera R. Guía para la configuración del Web Server Apache con Greenstone. ProdigioConsultores. [Sitio Web en Internet] 2010. Disponible en: <http://www.prodigioconsultores.org>

(7) Greenstone.org. Bibliotheca Digital Greenstone. Guía de Instalación. Capitulo 6.Personalizacion de su Instalación[Sitio Web en Internet].2007. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_personalizing.htm>

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

CAPÍTULO 14

14- Construcción de una colección con documentos PDF. Modificación de cadenas de formatos y macros

Lic. Elvia Vázquez Velázquez (elvia@cecam.sld.cu) Licenciada en Letras. Centro de Cibernética Aplicada a la Medicina (CECAM) I.S.C.M.H. Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad Médica de La Habana.

Palabras clave: Bibliotecas Digitales, Colecciones de documentos., Estrategias de Inclusión Digital, Greenstone, Sociedad de la Información, Software Libre.

14.1- Introducción

Entre las ventajas que justifican la aplicación del software Greenstone para crear colecciones digitales, se consideran :su fácil instalación con requerimientos mínimos de recursos en el sistema, y su libre disponibilidad de acceso que facilita el aprovechamiento especialmente en universidades, bibliotecas, y otras instituciones de servicio público de países sobre todo subdesarrollados (1), que lo conciben como recurso formidablemente útil para organizar, acceder y distribuir la información que poseen con diversos formatos.

Su flexibilidad admite modificar de forma creativa la apariencia individual y la posibilidad de personalizar las colecciones creadas en una biblioteca.

El modelo concebido en este acápite consiste en la construcción de una colección de archivos PDF con diferentes configuraciones en las cadenas de formato y la macro *extra.dm*

14.2- Modificación de cadenas de formatos y macros

La personalización de las colecciones permite a los bibliotecarios o interesados en el programa, efectuar cambios notables y modificar la apariencia que presentan originalmente, para hacerlas más atrayentes y funcionales a los usuarios finales. Esto se logra a través de las macros y de las cadenas de formato.

Dentro de la interfaz del bibliotecario de Greenstone, (GLI), en la opción Formato se despliega una pantalla con una cadena de caracteres en la que se combinan: funciones, componentes, variables con etiquetas HTML y etiquetas propias de Greenstone. Estos caracteres son metadatos que definen la configuración de las colecciones llamadas cadenas de formatos (2)

La configuración escogida en las cadenas de formatos se visualizarán posteriormente con los documentos en la página Acerca de ... de cada colección.

Una cadena de formato puede aplicarla a cualquier componente que se encuentre en la lista llamada "Elegir función". Ver Figura 14.1

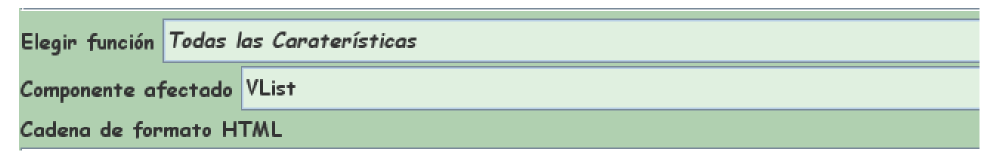


Figura 14.1. Componentes de opción Formato.

Incluye cada clasificador seleccionado durante el proceso de diseño y una lista de funciones determinadas previamente; algunas de estas requieren que se especifique una cadena de formato, para las cuales hay una lista descendente llamada "Componente afectado”. Ver Figura 14.2



Figura 14.2. Visualiza clasificadores y funciones seleccionados.

Las cadenas de formato que confeccione determinarán cambios con respecto a las que ofrece por defecto Greenstone, se distinguen rápidamente pues surten efecto dentro de la pestaña Formato; estas páginas no se hallan reunidas anteriormente en Greenstone, sino que se generan en el momento en que se configuran las cadenas de formato y son mostradas en pantalla cuando se visualizan las diferentes páginas concluyentes, esto no significa que haya que construir de nuevo la colección, sencillamente se actualiza y de inmediato se ven los cambios efectuados.

Greenstone dispone de valores predeterminados para cada cadena de formato; usted puede configurar cadenas nuevas que modifiquen el resultado de la apariencia de los documentos dentro de sus colecciones e invalidar las que genera por defecto el programa (2).

Dentro de los cambios previstos para el ejemplo a presentar en la construcción de la colección con documentos en formato pdf, cuenta además, desarrollar modificaciones a través de las macros; siempre presentes en el panel Formato mediante la opción Macros específicas de la colección.

Esta opción ofrece la posibilidad de configurar la colección al editar macros desde GLI, sin que para ello utilice los archivos de macros que se encuentran en el directorio de Greenstone.

14.3- Construcción de la colección con documentos en formato PDF

El ejercicio consiste en crear una colección con documentos en formato .pdf para lo cual se emplearon archivos de documentos originales de tesis de maestría defendidas en el Centro de Cibernética Aplicada a la Medicina (CECAM) y aunque la muestra es pequeña, resultan eficientes para lograr el objetivo previsto.

El primer paso es ejecutar el programa Greenstone por la opción GLI, para ello va al botón Inicio, todos los Programas, y elige Greenstone 2.82, seleccione *Librarian Interface Greenstone* (GLI). Inmediatamente aparece una consola con comandos MS-DOS y luego exhibe en pantalla la interfaz del bibliotecario con sus respectivos menús: Archivo y Editar

Una vez elegido el modo Avanzado de la opción Preferencias en el menú Archivo, el siguiente procedimiento es comenzar a crear una colección. Para ello de un clic en la opción Nueva dentro del menú Archivo .Ver Figura 14.3



Figura 14.3.Opción Nueva menú Archivo.

Este le solicita al usuario que debe llenar la información general acerca de la colección en los campos correspondientes: título de la colección una frase corta que refleja convenientemente el tema de la colección, en este caso la nombraremos "Tesis de maestría del CECAM", en la casilla Descripción de contenido se describe detalladamente la información que identifica la materia reunida en esta colección y que le aporta al usuario información suficiente para estimular el interés por ella. Véase Figura 14.4

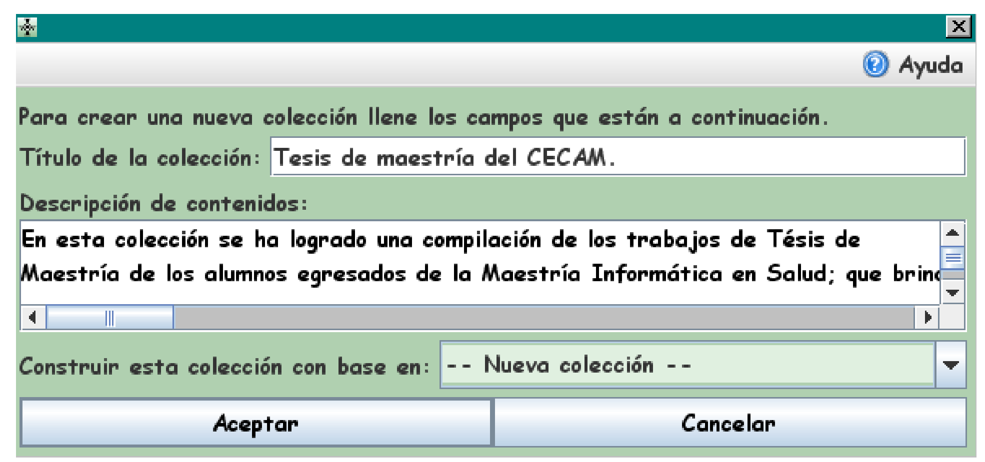


Figura 14.4.Opciones para crear una nueva colección.

Este texto es el que aparece posteriormente en la sección Acerca de en la página de esta colección.

En el parámetro Construir esta colección con base en seleccione Nueva colección tal como se percibe en la Figura 14.4 anterior

Pulse ahora sobre el botón Aceptar.

Todas las pestañas estarán activadas a partir de este momento. Realizados los procedimientos anteriores comienza a crearse la colección, presenta un mensaje que le indica al usuario este procedimiento. Figuras 14.5 y 14.6

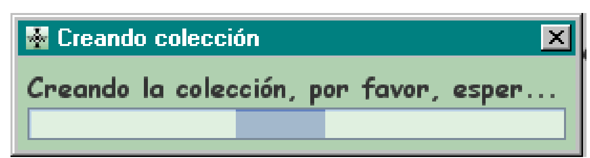


Figura 14.5 Mensaje al usuario para indicar que se crea la colección

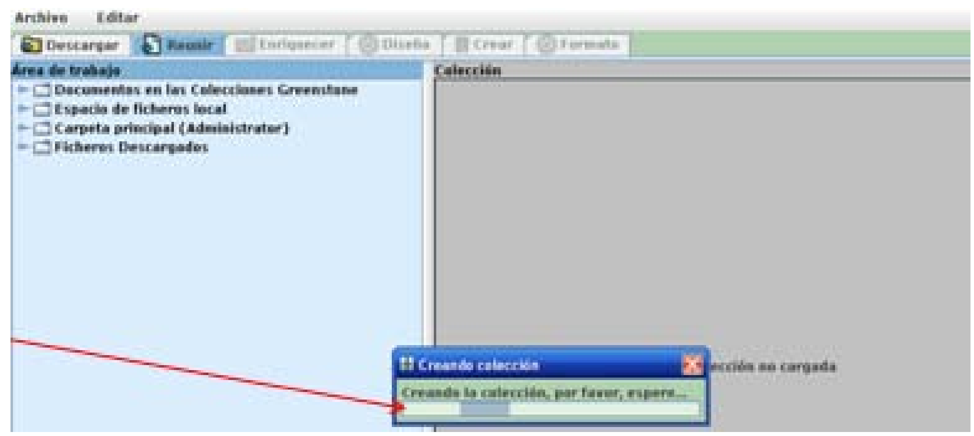


Figura 14.6 Proceso de carga de la colección

De inmediato se activa el espacio de trabajo de Greenstone pues asume el color verde en la pantalla del lado derecho, Ver Figura 14.7, el paso siguiente se apoya en las diferentes opciones disponibles en esta interfaz. La acción que se precisa a continuación es reunir los documentos previstos a incluir en la colección.

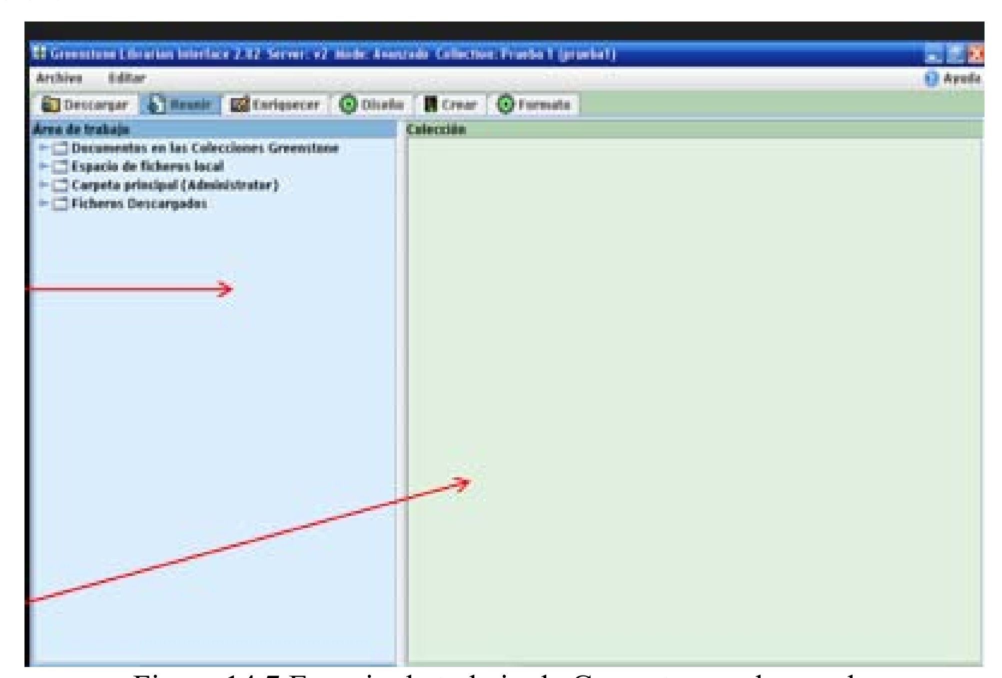


Figura 14.7 Espacio de trabajo de Greenstone color verde.

Obtención de los documentos a incluir en la colección.

Ahora la pestaña Reunir toma el color azul (activada), tiene dos áreas llamadas “Área de trabajo” y “Colección” situadas visiblemente debajo de las pestañas propias del software. Ver Figura 14.8

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

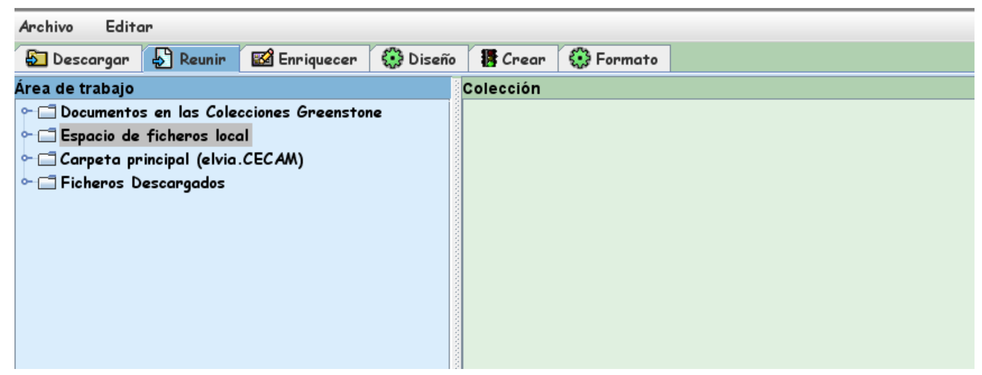


Figura 14.8 Pestaña Reunir con sus dos áreas.

El panel izquierdo o Área de trabajo'' permite explorar y localizar los documentos existentes en su computadora y que serán incluidos en la colección; este panel muestra una estructura jerárquica de las carpetas de trabajo equivalentes al explorador de Windows. Una vez que haya localizado los documentos, el próximo paso es pulsar sobre estos, y arrastrarlos al panel derecho.

Para trasladar los documentos hacia el espacio de Greenstone (color verde y derecho) el usuario selecciona los documentos (izquierda) (color azul), y con el mouse puede arrastrarlos y colocar los archivos en la parte derecha de la pantalla. Esta maniobra coloca los archivos convenientemente en la nueva colección, en este caso los documentos heredan metadatos propios del sistema de forma automática. Se copian sin desfigurar el sistema original de archivos existentes en su computadora.

Un aspecto importante es nombrar adecuadamente los archivos que formarán parte de la nueva colección, los títulos convienen que sean breves. Cuando se seleccionen, arrastren y coloquen grandes conjuntos de archivos en la nueva colección, la operación de copiado puede llevar tiempo según la cantidad de documentos que importe. La interfaz muestra en la parte inferior, el grado de avance en que se copian los archivos y el porcentaje de los que se han procesado.

14.4- Enriquecer los documentos con metadatos

El siguiente procedimiento a ejecutar consiste en enriquecer los documentos añadiéndoles metadatos, que son información organizada referente a los recursos de información, para describir los documentos que hayamos añadido en la pestaña Reunir. De esta manera se identifica cada recurso, y también sirve para recuperarlo posteriormente. La asignación de metadatos en Greenstone resulta decisiva y está asociada a la elección concreta de los índices de búsqueda y de clasificadores contemplados en las colecciones, así como otras facilidades, como ya hemos señalado previamente (3,4).

Este proceso se realiza en la pestaña Enriquecer, despliega una nueva pantalla de información la cual exhibe los directorios y subdirectorios de la colección con los documentos importados a la izquierda y a la derecha permite añadir los metadatos a los documentos individuales. Ver Figura 14.9.

El panel derecho de Enriquecer, expone el conjunto de metadatos seleccionados para los documentos. Los diferentes campos presentan un prefijo que comienza por dc--correspondiente a Dublin Core (5) considerado el más significativo ya que responde a la norma ISO generosamente aceptada como referente a nivel internacional--, dc aparece seguida del metadato que se trate (título, autor, descripción…etc.). El uso de metadatos ha devenido como una de las opciones más satisfactorias para organizar y recuperar posteriormente la información almacenada en las colecciones.

Para el ejemplo que proponemos se han decidido llenar de forma manual los metadatos de Título, Autor o Creador, Palabras claves, Descripción, Formato ,Fuente, Fecha. Ver Figura 14.9

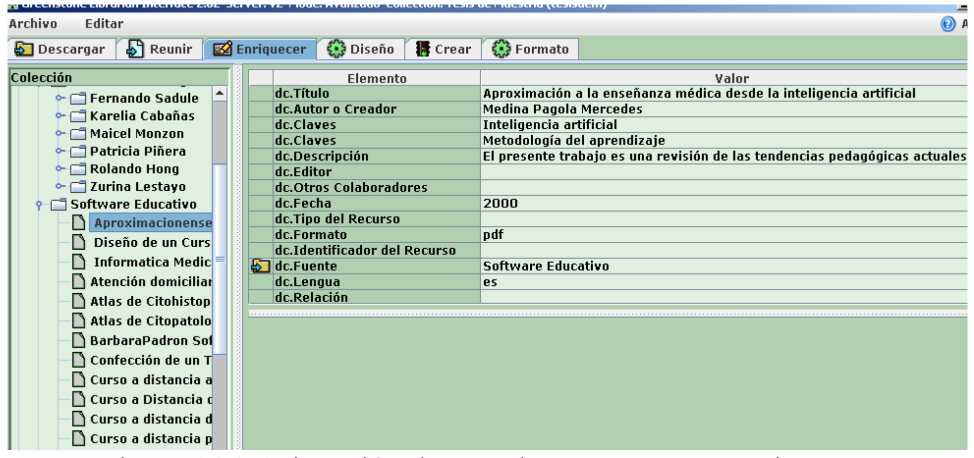


Figura 14.9 Asignación de metadatos en pestaña Enriquecer.

14.5- Diseño y configuración de la colección

Lograda esta vista se está en condiciones de configurar y establecer los mecanismos de búsquedas y navegación de la colección a través de la pestaña Diseño.

Esta pestaña le permite especificar la estructura, organización y presentación de la colección que se está creando. Este paso incluye una serie de pantallas separadas, cada una de las cuales trata un aspecto relativo al diseño de la colección y sirve, en efecto, como un equivalente gráfico al proceso normal de editar manualmente el archivo de configuración. (3).

En la Figura 14.10 a la izquierda, aparece una lista de las diferentes interfaces que el usuario puede configurar para lograr el diseño final, ellos son: Plugins de documento, Índices de búsqueda, Particionar índices, Clasificadores de exploración.

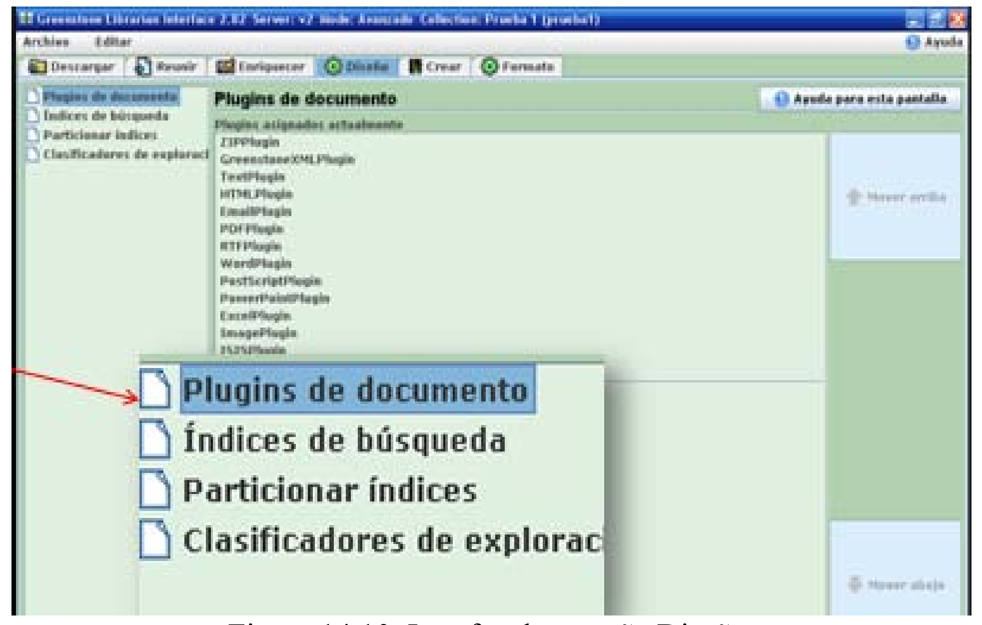


Figura 14.10. Interfaz de pestaña Diseño.

La forma que presentan y sus funciones varían entre ellas, por ejemplo, al dar un clic en el botón Plugins de documento, que aparece representado en la Figura 14.10 anterior, permite añadir,

eliminar o configurar los plugins o conectores originales así como cambiar el orden en el que se aplican a los documentos desde el lado derecho de la pantalla.

Recordemos que los plugins o conectores son pequeños programas que tratan documentos de variados formatos, html, word, pdf, etc. y los convierte en un formato de archivo Greenstone, estilo XML (6). Los plugins reconocen el formato de los documentos originales para procesarlos de igual forma durante el progreso de diseño y construcción de la colección para recuperarlos de manera correcta.

Puede aquí añadir, configurar o retirar plugins que se van a usar o no en la colección. No es obligatorio excluir plugins, pero si no resultan precisos algunos, puede apremiar un poco el procesamiento. Es importante conocer que la selección y configuración de los conectores desempeñan un papel básico durante el procesamiento de los documentos incluidos en las colecciones, si estos no son acertados algunos documentos pueden ser rechazados. Igualmente, el orden en que se asignan los conectores es fundamental para no ocasionar dificultades.

Como los documentos originales de esta colección están en formato PDF , se procede a eliminar de la lista aquellos que no son necesarios, excepto Greenstone XML Plugin y los que aparecen bajo la línea divisoria de esta pantalla, es decir Metadata XML Plugin, ArchivesInfPlugin, DirectoryPlugin; precisos para que funcione Greenstone. Ver Figura 14.11

Cuando los seleccione de la lista que aparece en el extremo inferior derecho al dar doble clic sobre él, aparece un cuadro de diálogo con varios argumentos que permiten especificar y refinar la configuración del plugin.

Para el caso que se ejemplifica elija el Plugins PDFPlugin, (ver Figura 14.11) recordar que los documentos están en formato pdf. Para ello debe seleccionarlo del extremo inferior derecho del listado que le ofrece esta interfaz.



Figura 14.11 Selección de plugin para documentos PDF.

14.6- Selección de los índices de búsqueda .Pestaña Diseño

14.6.1- Índices de Búsqueda o Consulta

Esta opción conforma la parte de los documentos que Greenstone dispone para que el usuario pueda recuperar información inclusive en volúmenes extensos posteriormente. Especifican cuáles criterios se emplearán para la búsqueda y están muy relacionados con los metadatos de acuerdo con las necesidades de los usuarios.

Preliminarmente Greenstone posee tres índices de consultas o búsqueda (Véase Figura 14.12): texto, título y fuente (Source), índices que contemplan las búsquedas o consultas a texto completo, por el metadato título y por el metadato fuente, heredados del documento original; sin embargo este último no satisface las necesidades demandadas por los usuarios .El índice de texto no debe descartarlo nunca pues determina las búsquedas a texto completo y debe disponerse en primer lugar siempre en la relación de índices asignados.

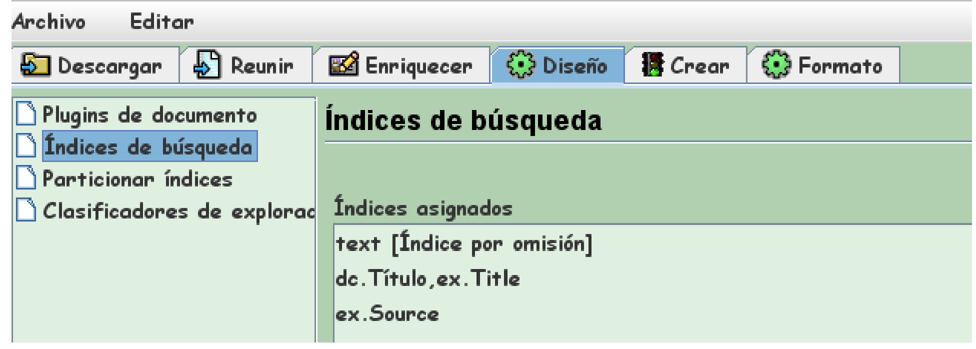


Figura 14.12 Disposición de los índices de búsqueda originales.

Para añadir uno nuevo debe dar un clic en la pestaña “Nuevo índice”. Ver Figura 14.13

Estos índices se seleccionan según el elemento de metadato expreso. Del mismo modo pueden configurarse índices que permitan consultar por varios campos a la vez y por textos completos en todos los documentos de una colección; esta es, ciertamente, una de las habilidades que proporciona Greenstone en todas las colecciones.

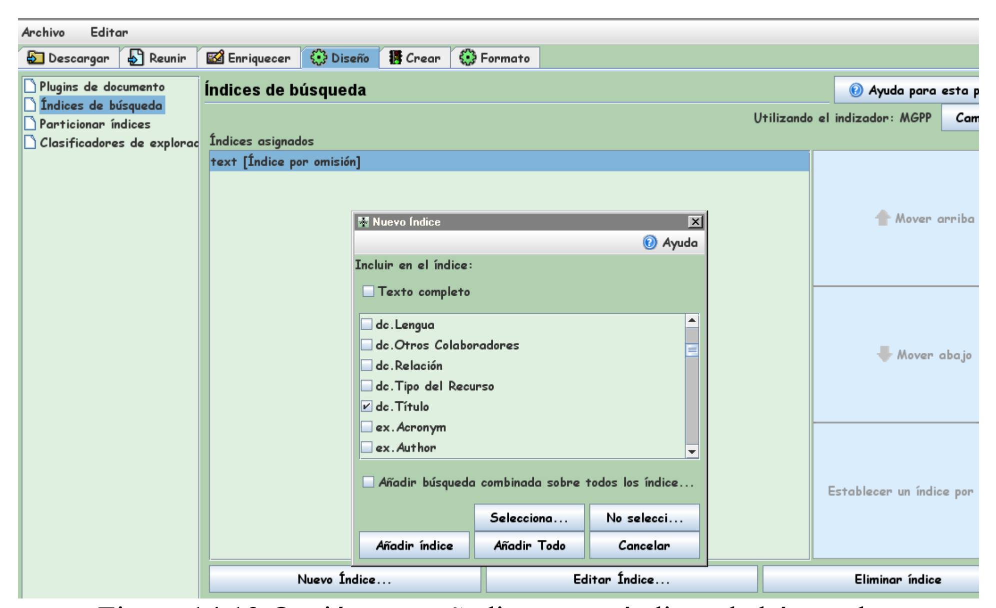


Figura 14.13 Opción para añadir nuevos índices de búsquedas.

Para el ejemplo se han modificado los índices originales y se añadieron índices de título, autor, palabras clave y fecha tal y como puede ver en la Figura 14.14 que a continuación se muestra.

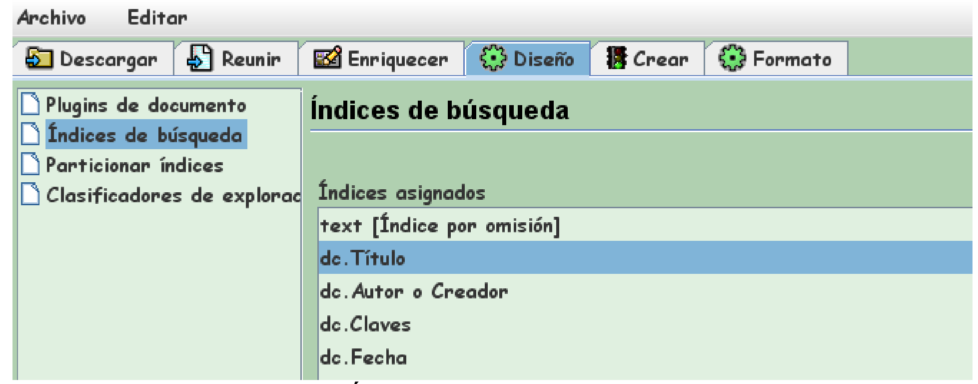


Figura 14.14 Índices de Búsqueda seleccionados.

El orden que tengan los índices en la lista de Índices asignados durante el proceso de configuración de la colección, será el mismo en que aparecerán en el menú de la página de búsqueda cuando usted despliegue en la casilla de Buscar en como se aprecia en la Figura 14.15.

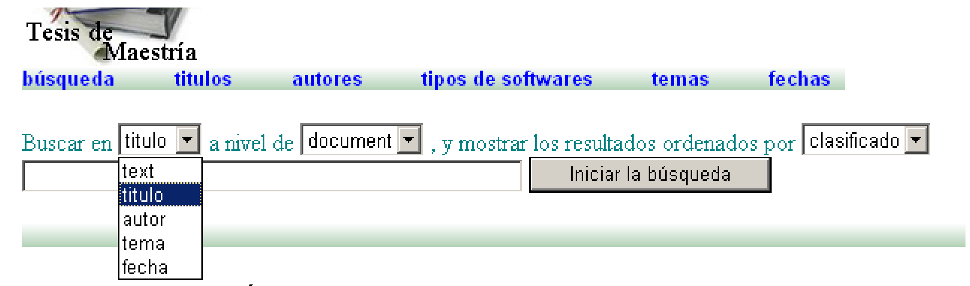


Figura 14.15 Índices de búsqueda según orden de configuración.

Aparecen en esta pantalla de "*Índices de búsqueda*" los indizadores disponibles en la colección, ellos son: MG, MGPP y Lucene potentes motores que facilitan la búsqueda de forma eficiente y de los que hemos explicado en el capitulo "*Sistema de Búsqueda de Información en Greenstone*".

14.7- Elección de clasificadores para establecer las estructuras de navegación según metadatos asignados.

14.7.1- Clasificadores de exploración

Corresponde entonces ahora activar la opción "Clasificadores de exploración" (Figura 14.16) esta es la forma de asignar "clasificadores", los que sirven para crear los índices de consulta de la colección, y se usan para navegar y buscar en ella.

Son las formas (botones) de mostrar los documentos de la colección al usuario, según los criterios de ordenación predefinidos.

Recordemos entonces que los clasificadores son pequeños programas escritos en lenguaje Perl (2). Ofrecen una ordenación y clasificación de todos los documentos. Se utilizan durante la creación de la colección organizándoles por título, autores, palabras clave, fechas, fuente…, etc. u otro metadato determinado durante la configuración y diseño de la colección; de ahí la importancia de estos últimos en la configuración de los clasificadores.



Figura 14.16 Opción Clasificadores de exploración.

Básicamente el sistema presenta dos clasificadores, los que puede anular para seleccionar nuevos del listado original que se despliegan en dicha pantalla. Ver Figura 14.17.

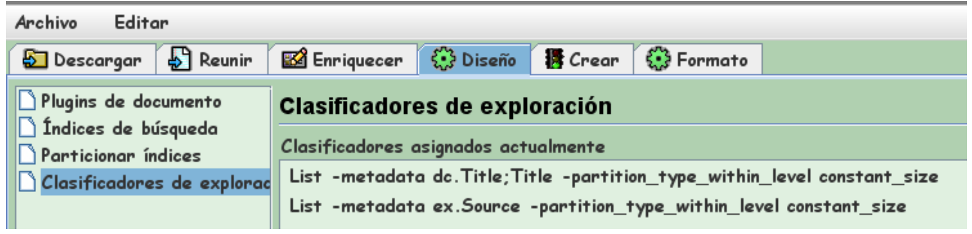


Figura 14.17.Clasificadores que presenta originalmente Greenstone.

Para quitar uno de la lista, selecciónelo y haga clic en "Eliminar el clasificador seleccionado". Ver Figura 14.18

Si desea añadir un clasificador nuevo, elíjalo por medio de la lista que ofrece la pantalla en el extremo inferior derecho y haga clic en "Añadir el clasificador especificado". Ver Figura 14.18

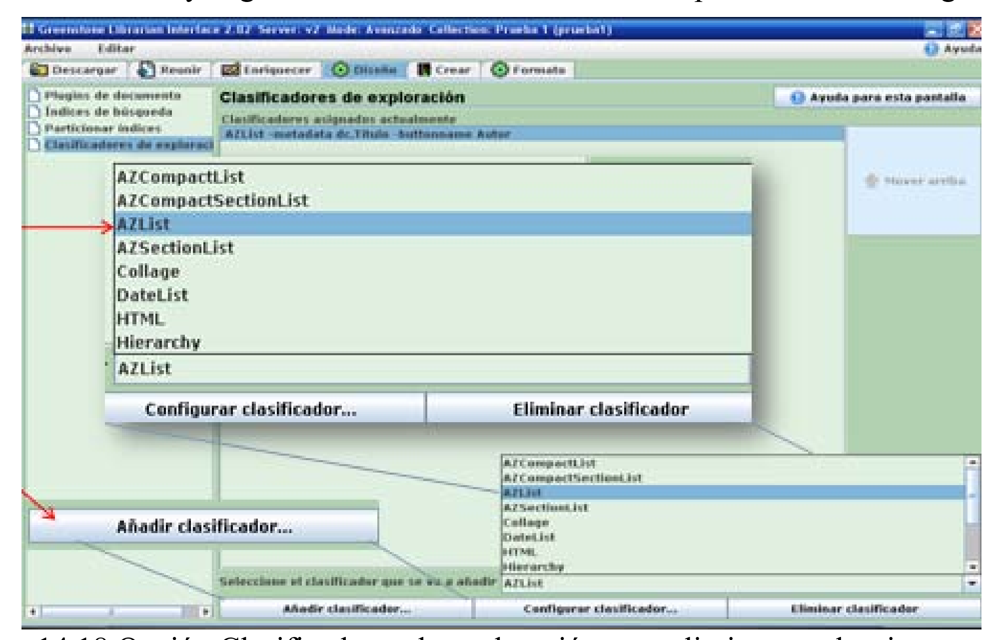


Figura 14.18 Opción Clasificadores de exploración para eliminar y seleccionar nuevos.

Aparecerá una ventana titulada "Argumentos de configuración".



Figura 14.19 Opción para configurar Clasificador de exploración AZList.

Las instrucciones para este cuadro de diálogo son las mismas que tienen los plugins. Una vez que haya configurado el nuevo clasificador, éste se añadirá al final de la lista de Clasificadores asignados actualmente. Ver Figura 14.20

Para la colección de muestra se han escogido el clasificador AZList –metadato Título, este crea una lista alfabética de títulos en el que reúne todos aquéllos que están dotados del campo de metadatos Título y los clasifica y divide por orden alfabético. Véase Figura 14.20.



Figura 14.20 .Selección y configuración de clasificadores según metadatos determinados.

Otro seleccionado es el clasificador AZCompactList para metadato autor o creador, palabras clave y tipo de software. Ver Figura 14.21.

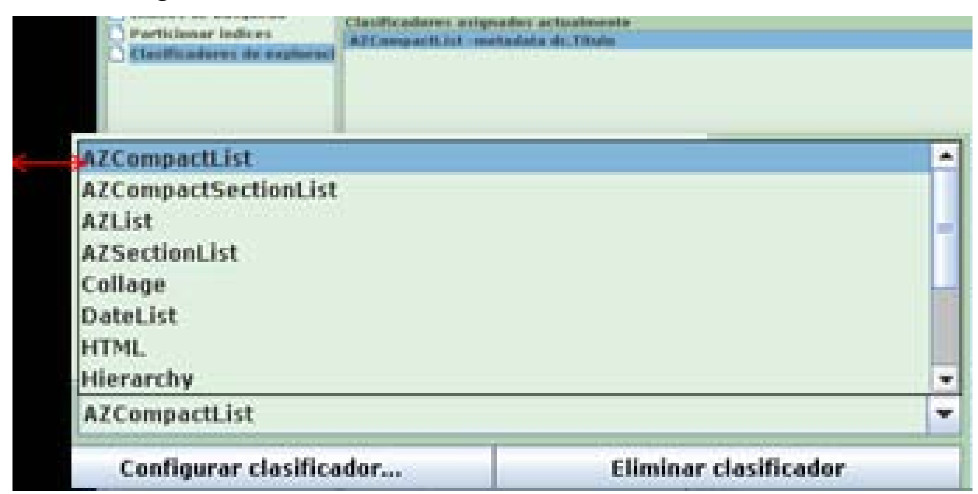


Figura 14.21 Selección de clasificador AZCompacList.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

Por último escogemos el clasificador DateList para el metadato fecha. Visible también en Figura 14.21 anterior.

En esta pestaña se define el orden de los clasificadores en la barra de navegación de la colección subordinada al gusto del creador de la colección. Para cambiarlo, seleccione el clasificador que desea mover y haga clic en "Mover arriba" o "Mover abajo", igual proceso que con los plugins.

Todos los clasificadores admiten el argumento buttonname, estos definen el contenido del botón de consulta en la colección y activan el clasificador determinado, básicamente es el nombre del argumento de metadato seleccionado ejemplo de esto lo puede ver en la Figura 14.22 ,al buttonname se le ha configurado el metadato fecha.

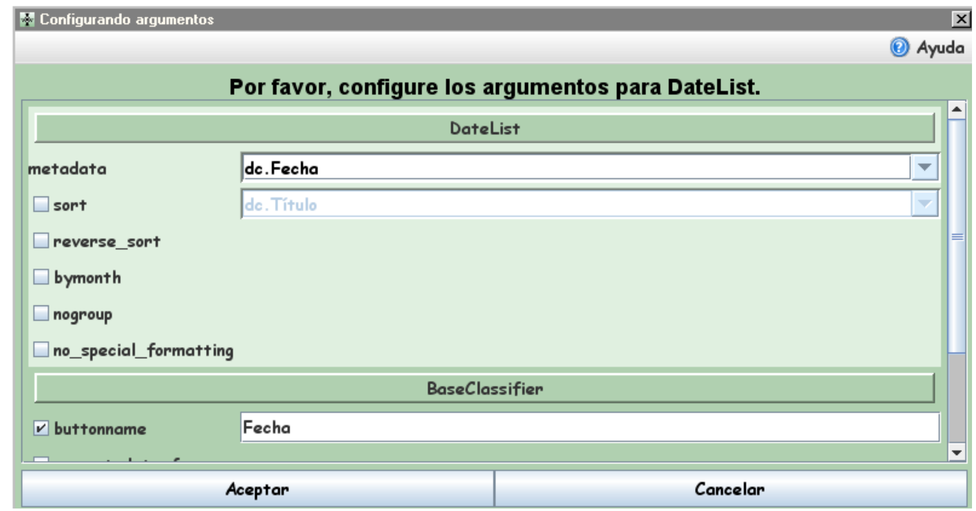


Figura 14.22 Selección y configuración clasificador DateList.

La barra de desplazamiento que aparece en la parte superior de la pantalla en las colecciones incluye la función búsqueda, siempre está disponible, seguida de botones convenientes a todos los clasificadores que se han definido, como lo ilustra la Figura 14.23 siguiente.



Figura 14.23 Botones o clasificadores y función Búsqueda en barra de desplazamiento en colección.

Todos los clasificadores generan una estructura jerárquica para mostrar un índice de consulta.

14.8- Configuración de la colección sobre la base de la pestaña Formato

14.8.1- Pestaña Formato. Opción General

La pestaña Formato en la opción General solicita datos para completar la información del documento según se precisen como son: el correo electrónico del administrador de la colección, el título de la colección, carpeta de la colección (nombre que asigna Greenstone automáticamente y no debe modificarse) los enlaces o URL del ícono de la página de la colección y de la principal solicitan la dirección proporcionada al ícono o imagen designada para la colección y para la página principal ,estas puede almacenarlas previamente en la carpeta images dentro de la carpeta Web del programa; al dar un clic en las opciones Explorar enlaza las imágenes allí guardadas (dentro de la carpeta Web) para identificar las páginas; la opción La colección será de acceso público define si la colección será pública o no (Ver Figura 14.24).

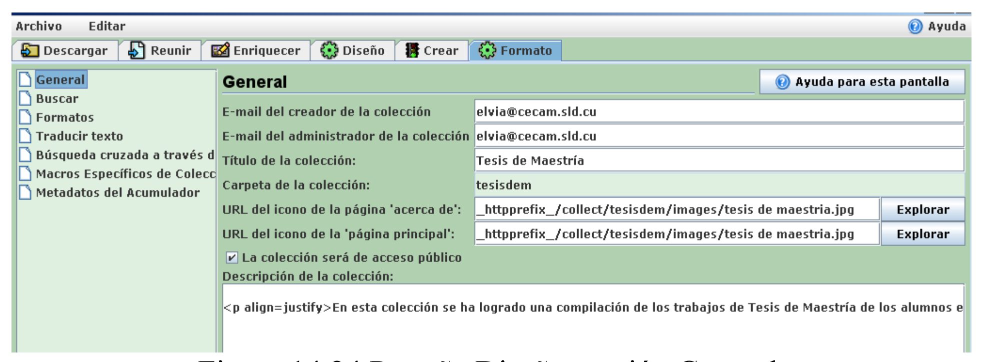


Figura 14.24 Pestaña Diseño, opción General.

Esta pantalla le permite además, en el espacio en blanco, modificar el texto de la descripción de la colección, que usted escribió cuando iniciaba el proceso de creación de la colección (menú Archivo (Nueva colección), si fuera necesario efectuar algún cambio. Se ha añadido además, la etiqueta <p align=justify> para justificar el texto de la colección de manera tal que la descripción de la colección en la página Acerca de se exponga al usuario con una apariencia más atractiva y personalizada.

14.8.2- Pestaña Formato. Opción Buscar

Esta opción está relacionada con los Índices de búsqueda de la pestaña Diseño.

En esta debe cambiar los datos que aparecen en inglés y editar los nombres en español para los índices de búsqueda (Figura 14.25) así cuando el usuario realice una búsqueda de información puede optar por el más favorable para él. Estos son los que se visualizan posteriormente en la colección en la función Búsqueda Ver Figura 14.25.

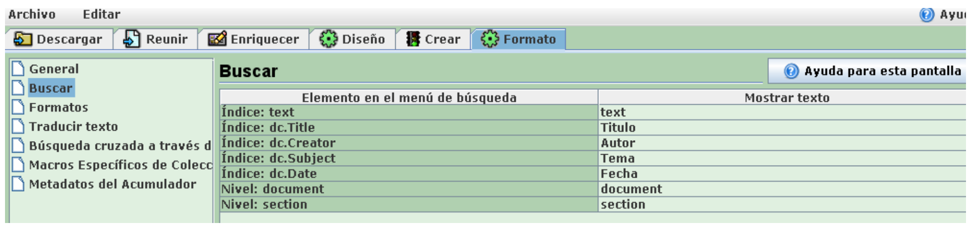


Figura 14.25 Modificación de los nombres de índices en inglés por nombres en español de acuerdo con las necesidades del usuario.

Para el usuario final quedaría como lo representa la Figura 14.26 siguiente.

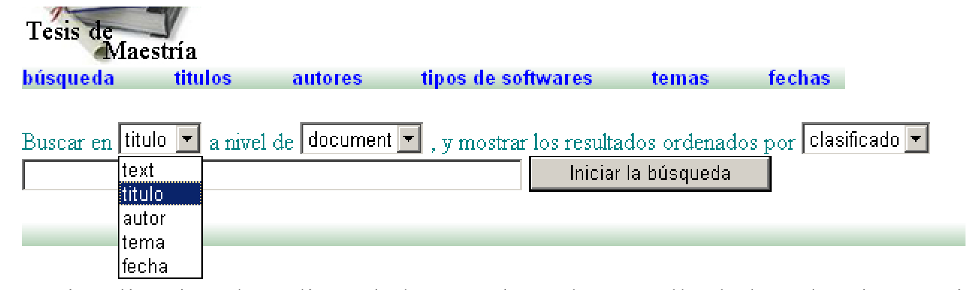


Figura 14.26 Visualización de índices de búsqueda en la pantalla de la colección opción Buscar.

14.9- Personalización de la colección mediante las cadenas de formato

14.9.1- Opción Formato

Anteriormente se explicó que esta opción incluye una pantalla con una cadena de caracteres en la

que se disponen: funciones, componentes, variables con etiquetas HTML y etiquetas propias de Greenstone. Y que son metadatos que definen la configuración de las colecciones conocidas como cadenas de formatos que establecen una tentativa de personalización de la colección; se modifican desde el panel Formato, opción Formatos. Ver Figura 14.27

Del mismo modo esta opción dispone de un listado superior que recoge los distintos formatos presentes en la colección, Hlist, VList,y DocumentButtons, los que determinan la presentación de los registros al navegar los Clasificadores de Exploración que se hayan escogido en la configuración de la colección. Document Heading, y Document Text son formatos que ofrecen el contenido completo de cada documento, según las opciones seleccionadas. Ver Figura 14.27

Igualmente aparece un menú desplegable en la parte inferior:

Elegir función, que sirve para determinar las posibilidades del formato,

El Componente afectado determina el formato de la lista al que se refiere.

Cadena de formato HTML, brinda el código y los detalles de cada formato. Ver Figura 14.27

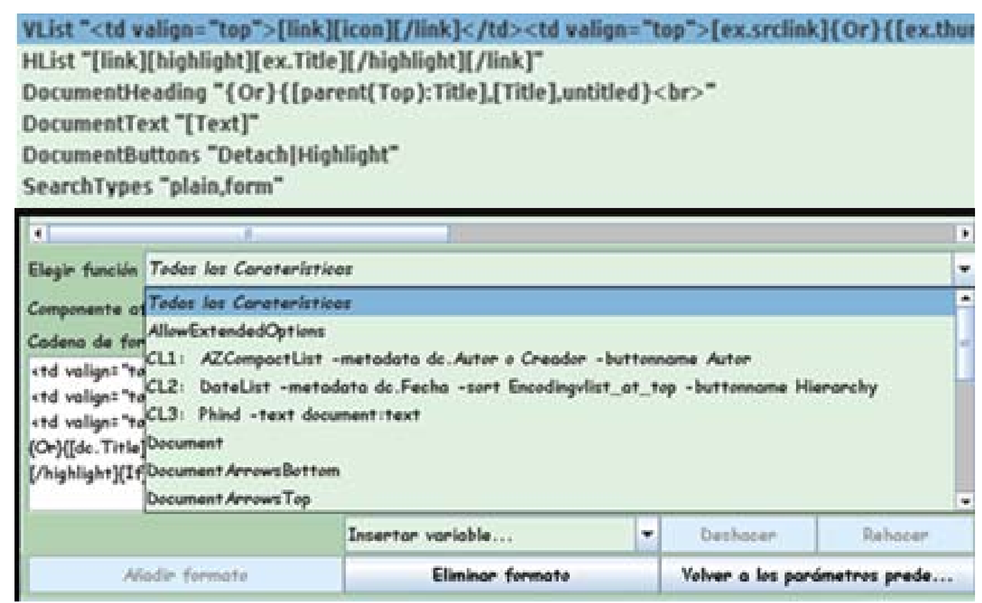


Figura 14.27 Opción Formato. Cadenas de Formato.

Se puede aplicar una cadena de formato a cualquier componente que se encuentre en la lista llamada "Elegir función", (Ver Figura 14.27) la cual incluye cada clasificador seleccionado y una lista de funciones determinadas previamente. Para cambiar una cadena de formato, selecciónelo de la lista, modifique sus parámetros y haga clic en "Añadir formato".

El primer cambio previsto consiste en modificar la cadena de formato original del clasificador CL1, para ello se elimina la primera línea:

<td valign="top"> [link][icon][/link]</td>

Cuyo efecto crea un enlace al ícono de la versión html del documento, no necesaria para el ejemplo con documentos PDF.

Elegimos el mecanísmo:

[ex.srclink ex.srcicon]}[ex./srclink]

Que crea un enlace etiquetado por el ícon estándar de PDF.



La intención es acceder a los documentos mediante su versión original, utilizaremos entonces la siguiente cadena de formatos:

Para el clasificador Cl1: AZList-metadata dc.titulo-buttonname título Componente afectado VList

<td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
[/highlight]{If}{[ex.Source],
<i> ([ex.Source])</i>}</td>

El resultado será como muestra la Figura 14.28.

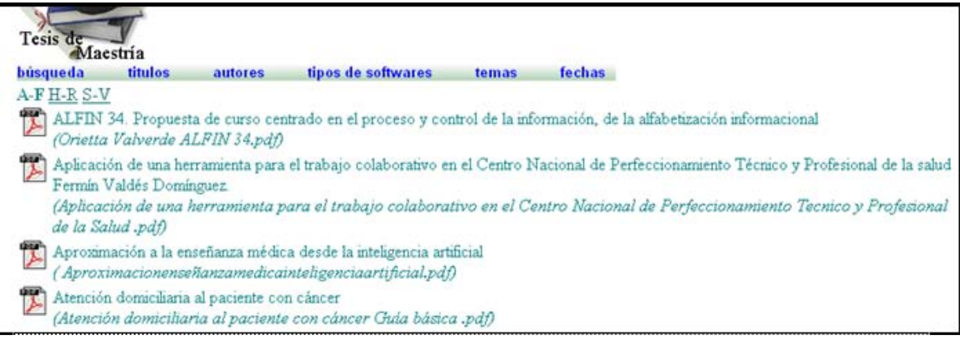


Figura 14.28. Resultado de la configuración de cadena de formato para el clasificador Cl1: AZList-metadata dc.titulo-buttonname titulo Componente afectado Vlist.

Como puede apreciar se listan los documentos por títulos ordenados alfabéticamente, esto lo confirman las letras A-F H-R S-V en la parte superior del listado, rangos en que se encuentran los documentos de tesis organizados según las letras mencionadas. Aparece además entre paréntesis ( ) el título original del documento sin procesar lo que no resulta apropiado para una colección con documentos con las características de textos típicos como son los de tesis de maestría o trabajos de investigación realizados con la finalidad de obtener un título

Por eso se elimina la última línea que hace referencia al título del archivo fuente u original [/highlight]{If}{[ex.Source],
<i> ([ex.Source])</i>}</td>, esto regulariza espacio disponible y concede un aspecto más elegante a cada página concluyente. Con la siguiente cadena de formato

```
<td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
```

Y el resultado será como muestra la Figura 14.29

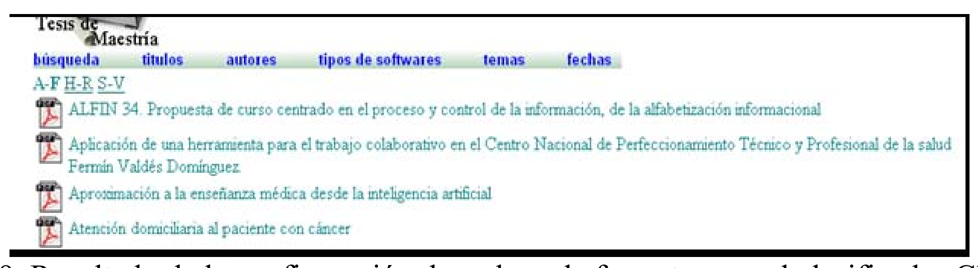


Figura 14.29. Resultado de la configuración de cadena de formato para el clasificador Cl 1: AZList-metadata dc.titulo- buttonname titulo. Componente afectado VList sin la última instrucción.

Si desea insertar variables de Título, Autor y las etiquetas <b><u> para este mismo clasificador con la siguiente cadena de formato podría quedar tal y como muestra la Figura 14.30.

```
<td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
<b><u>{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}</b></u>
[/highlight]
{If}{[dc.Creator],<br> Autor: [dc.Creator]</i>}
{If}{[dc.Date],<br> Fecha: [format:dc.Date]</i>}
<br>
</td>
```

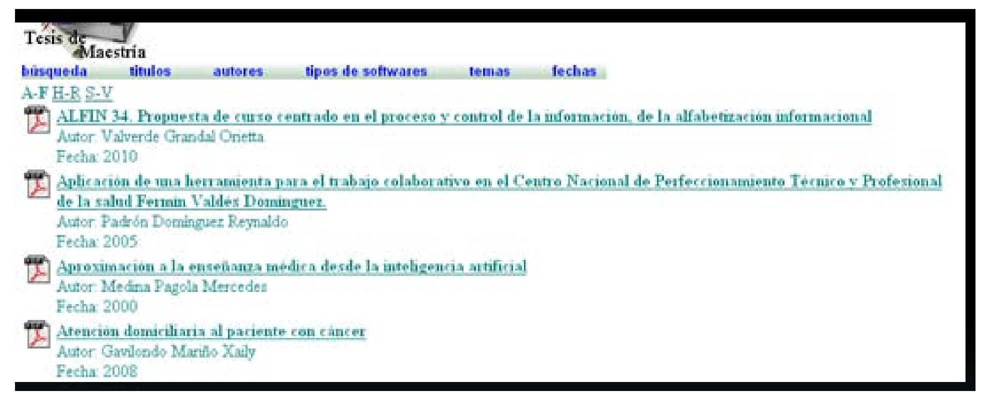


Figura 14.30. Resultado de la configuración de cadena de formato para el clasificador Cl1: AZList-metadata dc.titulo-buttonname título Componente afectado VList con variables Título, Autor y Fecha.

Se logra que el título aparezca en letra negrita y subrayada; pero es el ícono de PDF el vínculo al documento.

Si además le agrega la variable [dc.Description] con el fundamento {Or} delante

```
<td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
<b><u>{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}</b></u>
[/highlight]
{If}{[dc.Creator],<br> Autor: [dc.Creator]</i>}
{If}{[dc.Date],<br> Fecha: [format:dc.Date]</i>}
<br>
{Or}{[dc.Description]<b>dc.Descripción</b>}
<td align="left">[highlight] [srclink]
</br>
</td>
```

Se visualiza como la siguiente Figura 14.31.

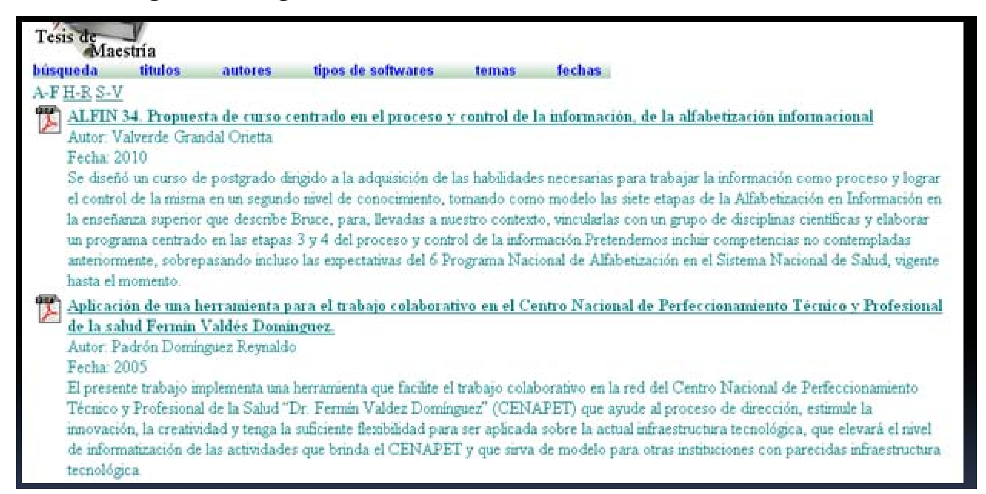


Figura 14.31 Resultado de la configuración de cadena de formato para el clasificador Cl1: AZList-metadata dc.titulo-buttonname titulo Componente afectado VList con variable Descripción.

Esta cadena de formato con la inclusión de la variable [dc.Description] beneficia al usuario con una idea aproximada, a simple vista, del contenido de ese documento sin tener que acceder al texto

completo; esto está fundamentado por el metadato dc.Descripción seleccionado en la pestaña Enriquecer en el momento de la configuración.

Como se ha comprobado a partir de la Figura 14.28 y hasta la Figura 14.31 con el mismo clasificador AZList metadato título se insertan diferentes variables en la cadena de formato, obteniéndose cambios que se manifiestan en el momento de visualizar la página de la colección.

Para mostrar los autores se utilizó el clasificador AZCompactList que crea un compactado de cada criterio, en este caso de los autores.

Para el clasificador Cl2: AZCompactList-metadata dc.Creator-buttonname Autor Componente afectado VList

Con la siguiente cadena de formato:

```
<td valign=top>{If}{[srcicon],[srclink][srcicon][/srclink], [link][icon][/link]}</td>
<td valign="top">[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
</td>
```

Queda como muestra la Figura 14.32.

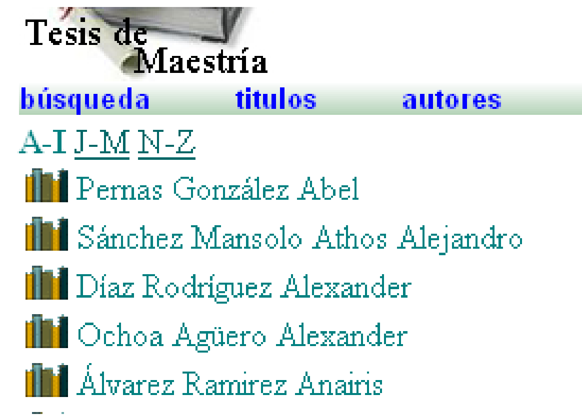


Figura 14.32. Resultado de la configuración de cadena de formato para el clasificador Cl2: AZCompactList-metadata dc.Cretor-buttonname Autor Componente afectado Vlist.

Este clasificador  genera estanterías que almacenan la cantidad de documentos por autor. Es decir que en cada ícono se almacenan los autores según los títulos que incluyan cada uno.

Si al final de la cadena anterior agregamos la instrucción [numleafdocs] ,tiene el efecto de especificar entre paréntesis el número de elementos que se encuentran dentro de un estante, en el ejemplo anterior presenta la cantidad de títulos que almacena cada autor. Ver Figura 14.33

Para contar la cantidad de documentos agrupados

```
<td> {If}{[numleafdocs],<i> ([numleafdocs])</i></small>}</td>
```

Numleafdocs: No será establecido para los documentos.

Tesis de Maestría

búsqueda titulos autores tipos de softwares

A-I J-M N-Z

Pernas González Abel (1)

Sánchez Mansolo Athos Alejandro (1)

Díaz Rodríguez Alexander (1)

Ochoa Agüero Alexander (1)

Álvarez Ramirez Anairis (1)

Barrios Blanco Abel , Reyes Candia Valia (1)

Figura 14.33. Resultado de la configuración de cadena de formato para el clasificador Cl2: AZCompactList-metadata dc.Cretor-buttonname Autor con la instrucción Numleafdocs.

Si da un clic en el ícono del estante del autor Pernas González Abel, u otro autor, se despliega 1 documento como bien le indica el número 1 entre paréntesis (1) destacado en la Figura 14.34.



Figura 14.34. Desglose de un estante con autor y cantidad de documentos.

Para esta colección también se configuró el clasificador Temas con las palabras claves que contenían los documentos procesados, y también se escogió un clasificador AZCompactList con la siguiente cadena. Vea el resultado en la siguiente Figura 14.35.

```
<td valign=top>{If}{[srcicon],[srclink][srcicon][/srclink],
[link][icon][/link]}</td>
<td valign="top">[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
</td>
```

Tesis de Maestría

búsqueda titulos autores tipos de softwares temas

A-D E-H I-N P-R S-W

Alfabetización en Información

ALFIN 34

Alimentación y nutrición humanas

Anatomía

Figura 14.35. Resultado de la configuración de cadena de formato para el clasificador Cl4: AZCompactList–metadata dc.Claves –buttonname Temas Componente afectado Vlist.

Esta estantería agrupa tesis que presenten las palabras clave mostradas en pantalla,por ejemplo la palabra Anatomía,puede incluir varios documentos que se relacionen con esta temática.

Para el clasificador Cl3: DateList –metadata dc.date –buttonname Fecha cuya cadena de formato es la siguiente:

```
<tdvalign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}
[/highlight]
</i>
```

Este clasificador DateList genera una lista de selección de intervalos de fechas como lo presenta la Figura 14.36.

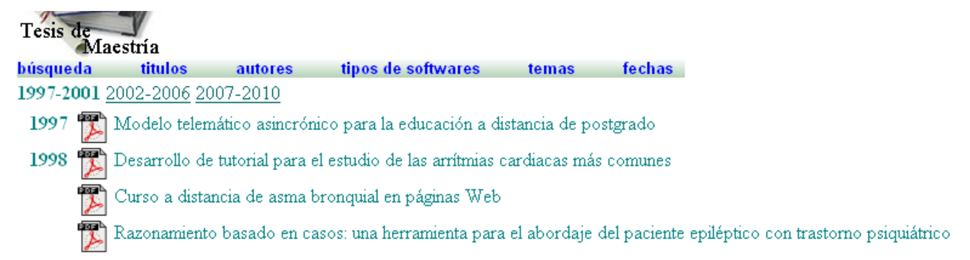


Figura 14.36. Resultado de la configuración de cadena de formato para el clasificador Cl 3: DateList –metadata dc.date –buttonname Fecha Componente afectado Vlist.

Es importante recordar, que todos estos cambios en las cadenas de formato, se vizualizan inmediatmante sin que para ello tenga que crear de nuevo la colección, lo que soluciona inmediatamente,sin ofrecer considerable tiempo en el trabajo,la apariencia final de la colección de forma apremiada.

Estos ejemplos de cadenas de formatos demuestran que con pequeñas modificaciones y el uso de diferentes variables se pueden lograr cambios en los resultados finales de visualización de los documentos que conforman las diferentes colecciones digitales.

Se acentúa en el cuidado y esmero a tener cuando diseñe las cadenas de formatos; una simple omisión de una etiqueta, una llave, instrucción o signo utilizado de forma incorrecta afecta, indiscutiblemente, en la apariencia de la página o no logra el efecto que concibió originalmente en su diseño.

14.10- Personalización de la colección mediante modificación de macros

Existen otras posibilidades para personalizar la apariencia de la colección en lo que a presentación de la página se refiere. Tendremos en cuenta, para este caso, la opción Macros específicos de la colección, aludidas al principio de este capítulo.

Para trabajar con esta opción, se procedió de la siguiente manera.

Se activó la opción Macros específicos de la colección, situada al lado derecho de la opción Formato, Ver Figura 14.37

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0



Figura 14.37. Opción Macros específicas de la colección.

Las macros están incluidas en un archivo llamado *extra.dm* (7)

Esta sección admite editar el archivo que corresponde a la macro extra.dm de esta colección y se ejecuta desde GLI.

Se realizaron algunos cambios que se presentan posteriormente en la página en la sección Acerca de. Primeramente se procedió a cambiar el título de la sección Acerca de de esta colección para ello se escribió debajo de las macros

extra.dm file y _textabout_ el siguiente texto “Colección Tesis de maestría del CECAM”

```
package about
_textabout_ {
<h3>Colección Tesis de maestría del CECAM.</h3>
_Global:collectionextra_
</div>
}
```

La macro _textabout_ permite suplantar el texto que aparece en Acerca de por otro que desee, para el ejemplo se decidió “Colección Tesis de maestría del CECAM” por ser el título de la colección, y el que permite posteriormente al usuario identificar la colección que explora.

textabout se escribe entre las etiquetas <h3> </h3> que definen el tamaño de la letra de este texto.

Además se le añadió otro texto sucesivo al anterior y separándoles con la etiqueta
 para dar cierto espacio entre un texto y otro, el que precisa la descripción de esta colección de la siguiente forma

<h4>Información descriptiva sobre esta colección</h4>

Por tanto el puntaje de este argumento será más pequeño que el anterior para establecer diferencias entre uno y otro mensaje.

Otras transformaciones realizadas fueron adicionarle un pie de página a cada página, para ello se utilizó la macro _footer, al final del espacio de texto. El contenido quedó al centro pues se emplearon las etiquetas <center></center> y <small> y el texto aparece en un tamaño más pequeño que el resto de la página, la etiqueta <p> </p> para que el pie de página quedara en otro párrafo aparte.También al lado de la línea _pagefooterextra_ se le agregó

para dar un margen al texto que dice Colección creada por Elvia Vázquez Velázquez Copyright 2011 Biblioteca Virtual CECAM

Package about

footer_ {
pagefooterextra

<center> Colección creada por Elvia Vázquez Velázquez<p><center>
<small>Copyright 2011 Biblioteca Virtual CECAM</p></small></center>}

El resto de las macros que aparecen en esta opción se mantienen intactas. De tal manera quedaría como lo muestra la siguiente Figura 14.38

```
# extra.dm file
package about
_textabout_ {

<h3>Colección  Tesis de maestría del CECAM</h2><br><h4>Información descriptiva sobre esta colección</h4>
_Global:collectionextra_
</div>
}
# You can add collection specific macros in here
package about
# Lines starting with a '#' are comments.
# Remember to include the package declaration
package Style
_footer_ {
_pagefooterextra_ <br><br><center>Colección generada por  Elvia Vázquez Velázque:
<center><small>Copyright 2010. Biblioteca Virtual CECAM</small></center>
}
# You can add collection specific macros in here
# Lines starting with a '#' are comments.
# Remember to include the package declaration
```

Figura 14.38. Modificación de macros desde opción Macros específicos de la colección.

La macro _footer_ , proporcionó el mismo resultado en todas las páginas de la colección.

La Figura 14.39 presenta el resultado de las modificaciones expresadas anteriormente y que se aprecian en la página de esta colección de “Tesis de maestrías del CECAM”.La figura incluye además todos los cambios establecidos durante el diseño y configuración de esta colección referentes a clasificadores o botones de navegación,título,descripción,el ícono que la identifica, la opción de buscar en la colección por el índice de texto , y pie de página incorporado.Ver Figura 14.39

Estas transformaciones se distinguieron rápidamente con sólo un clic en el botón actualizar y a través de la opción Vista previa de la colección

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

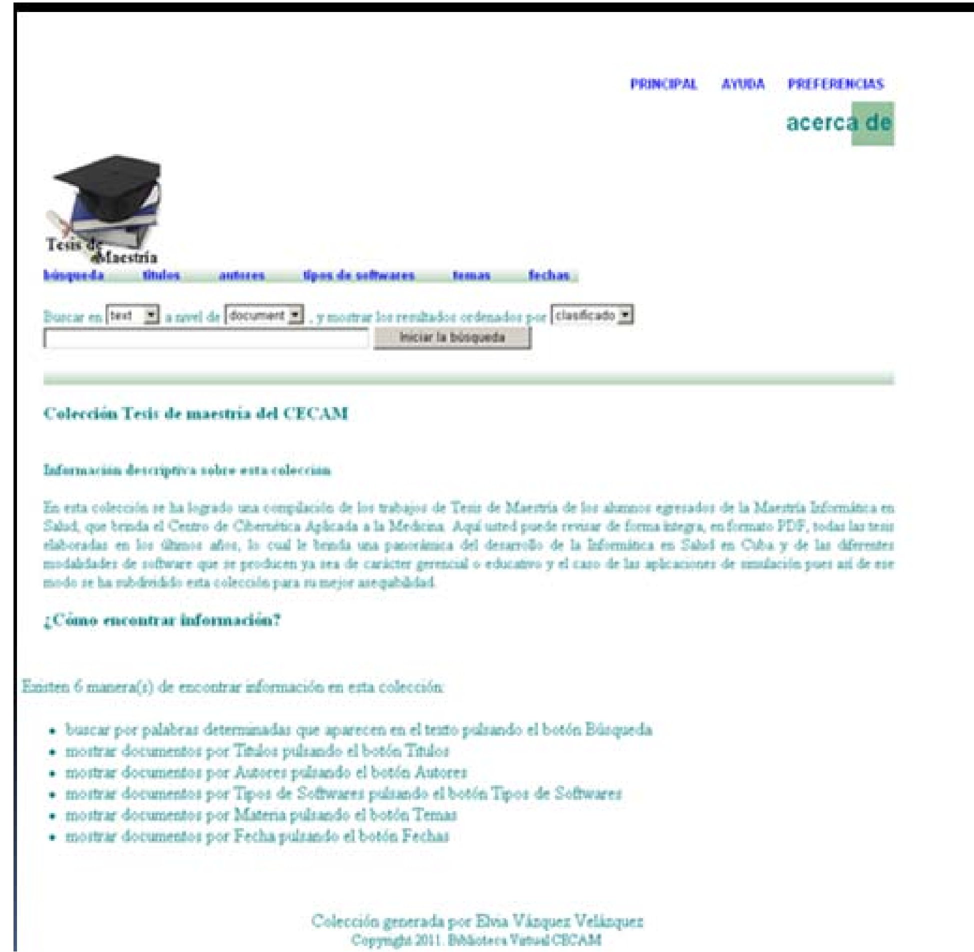


Figura 14.39 Página principal de la colección. Efecto logrado con cambios en la opción Macros específicos de la colección.

Con los ejemplos mostrados se ha pretendido configurar y personalizar la apariencia de una colección Greenstone desde dos maneras posibles para lograr páginas nuevas que distan de las generadas por el propio programa Greenstone.

14.11- Consideraciones finales

Con las diferentes cadenas de formatos ejemplificadas en este trabajo podemos concluir que es posible y diversa la combinación de formas con las que pueden configurarse estos comandos de formatos, sólo es necesario que se hagan diferentes pruebas hasta lograr la deseada y resulte más agradable a la vista, la página de cada una de las colecciones que conforman su biblioteca digital.

La flexibilidad ofrecida por Greenstone, tanto para cadenas de formato y macros, resulta fundamental para lograr una interfaz con una apariencia en la que se exprese la creatividad del bibliotecario preocupado por un usuario final que se beneficie con todo el potencial brindado por la colección consultada y alcance a resolver los problemas esenciales de recuperación de información textual.

14.12- Referencias Bibliográficas

(1) Witten I H, Spano D. Greenstone en América Latina: Una ola expansiva de más de 11.000 km[Libro en Intenet] En: Greenstone: Un software libre de código abierto para la construcción de bibliotecas digitales.2010, UNESCO. Montevideo. Disponible

en<http://www.unesco.org.uy/ci/fileadmin/comunicacion-informacion/Greenstone-paraweb.pdf>

(2) Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 2. Explotación Óptima de sus Documentos. [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

(3) Bainbridge D, McKay D, Witten H. I. Guía del programador. Biblioteca digital. Greenstone. [Sitio Web en Internet].Departamento de Informática. Universidad de Waikato, Nueva Zelanda. Marzo de 2004.Disponible en:<http://www.greenstone.org/manuals/gsdl2/es/html/Develop_es_all.html>

(4) Wiki Greenstone.org. Ejercicio tutorial de Greenstone. Editando set de Metadatos. [Sitio Web en Internet] 2010. Disponible en:<http://wiki.greenstone.org/wiki/gsdoc/tutorial/en/index.html>

(5) Dublin Core Metadate Iniciative. Conjunto de Elementos de Metadatos Dublin Core, Version 1.1: Descripción de Referencia.[Sitio Web en Internet]. 2003-06-02. Disponible en <http://es.dublincore.org/documents/2003/06/02/dces/index.shtml.rdf>

(6) Masternagazine. Definición de Plug-in.[Sitio Web en Internet] 16 de Febrero de 2005. Disponible en: <http://www.mastermagazine.info>

(7) Wiki Greenstone.org. All about macros.[Website in Internet] 3 November 2006. Available from: < http://wiki.greenstone.org/wiki/index.php >

CAPÍTULO 15

15- Creación de una colección de imágenes

Lic. Elvia Vázquez Velázquez (elvia@cecam.sld.cu) Licenciada en Letras. Centro de Cibernética Aplicada a la Medicina (CECAM) I.S.C.M.H. Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354

Palabras clave: Bibliotecas Digitales, Colecciones de documentos., Economía Digital,Estrategias de Inclusión Digital, Greenstone, Sociedad de la Información, Software Libre, UNESCO.

15.1- Introducción

Greenstone se compone de distintas herramientas para desarrollar colecciones digitales de manera fácil e independiente con todo tipo de materiales y desarrollar el acceso universal a la cultura. Esta es su razón de ser. (1)

Proporciona el acceso al contenido de estas colecciones en formatos diferentes y atractivos, inclusive con animación, formatos gráficos, información, etc., para adaptarse a las necesidades de usuarios, uno de ellos son las imágenes.

Las imágenes digitales son instantáneas electrónicas tomadas de una circunstancia o escenario y pueden ser escaneadas de documentos como textos impresos.

Cuando se insertan imágenes en un documento electrónico, es necesario considerar la calidad de la imagen y su tamaño en Kb, si se usan imágenes grandes, la descarga del documento tomaría demasiado tiempo igualmente la calidad sería más alta, determinada principalmente por su resolución, la cual permite distinguir hasta el más pequeño detalle.

Los formatos comunes de imágenes son JPEG/JPG (*Joint Photographic Experts Group*) es un algoritmo diseñado para comprimir imágenes con 24 bits de profundidad o en escala de grises. JPEG es también el formato de fichero que utiliza este algoritmo para comprimir imágenes, el formato de archivos JPEG se abrevia frecuentemente JPG debido a que algunos sistemas operativos sólo aceptan tres letras de extensión (2), considerado un buen formato para mostrar fotografías en un formato compatible con la Web, pues acepta millones de colores lográndose comprimir bastante.

15.2- Características de MARC XML

Se trata de empaquetar los documentos MARC en u estándar más flexible y comprensible. El lenguaje XML trabaja con etiquetas que engloban la información y de esta manera los usuarios pueden trabajar con esos datos de manera más adecuada a sus necesidades (2).

15.3- Selección de imágenes a incluir en la selección

Para crear una colección de imágenes hemos previsto un ejercicio con imágenes extraídas de un documento descargado de la Web, titulado “La terapia floral de Bach” de la Lic. Martha Cruz Torres.

Independientemente del sugerente y atrayente tema en el que se describen las características del sistema floral de Bach, hemos considerado apreciable el uso de las imágenes representadas en este artículo.

De las 38 imágenes de flores insertadas en el artículo hemos escogido al azar una muestra pequeña para desarrollar el ejercicio concebido y lograr el objetivo inicial: crear una pequeña colección de imágenes.

Estas han sido modificadas y personalizadas con SnagIt, innovador y útil programa para realizar capturas de pantalla y editarlas a nuestro gusto, compuesta por un variado conjunto de herramientas destinadas a la correcta edición de dichas imágenes (3); con este se les ha asignado un borde a cada una con un color similar al de las flores así como también se les ha proporcionado el mismo tamaño a cada una en aras de ganar similitud en cada imagen. (Ver Figura 15.1)



Figura 15.1.Selección de imágenes y personalizadas con SnagIt.

15.4- Construcción de una colección pequeña de imágenes

Para comenzar a trabajar deben localizar en sus computadoras la interfaz del bibliotecario de Greenstone desde el botón Inicio, Todos los programas, Greenstone versión 2.82 y dar un clic en el icono *Librarian Interface* (GLI). (Ver Figura 15.2)

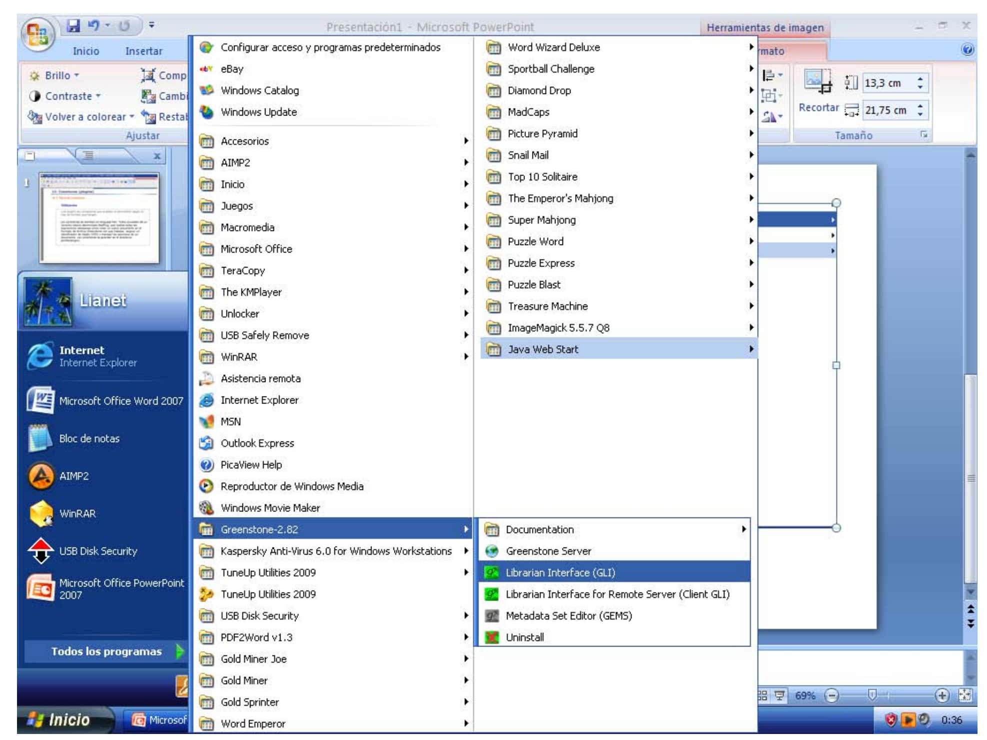


Figura 15.2. Ruta donde localizar GLI.

Una vez abierta la interfaz del bibliotecario, comenzamos a crear la colección de imágenes, para

ello en el menú Archivo seleccionar la opción Nueva, a continuación llene los campos con la información solicitada y de forma adecuada, en este caso el título de la colección es “*Flores de Bach, imágenes*” y la descripción de la colección que define el tema que trata la colección tal y como se muestra en pantalla. Para la opción Base esta colección en, seleccione Nueva colección. Véase Figura 15.3

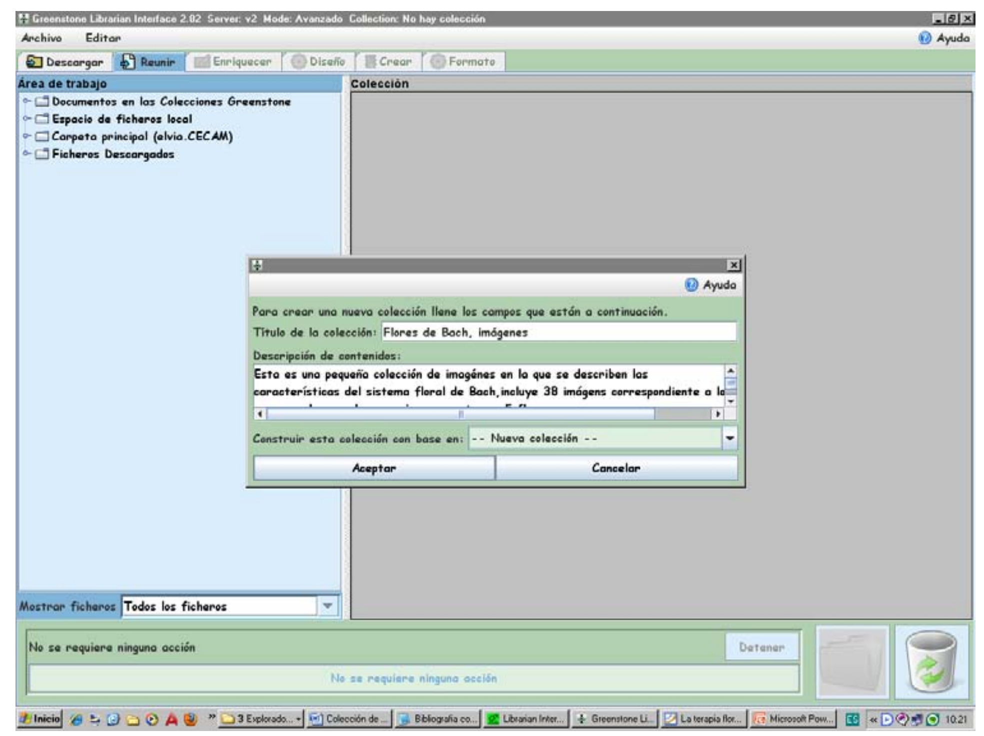


Figura 15.3 Definición de nueva colección.

Comienza el proceso de creación de la colección y aparece en pantalla el cuadro de diálogo (Ver Figura 15.4) que le informa al usuario que espere unos minutos para comenzar a localizar los documentos en su computadora a través de la pestaña Reunir tal como aparece en la Figura 15.4

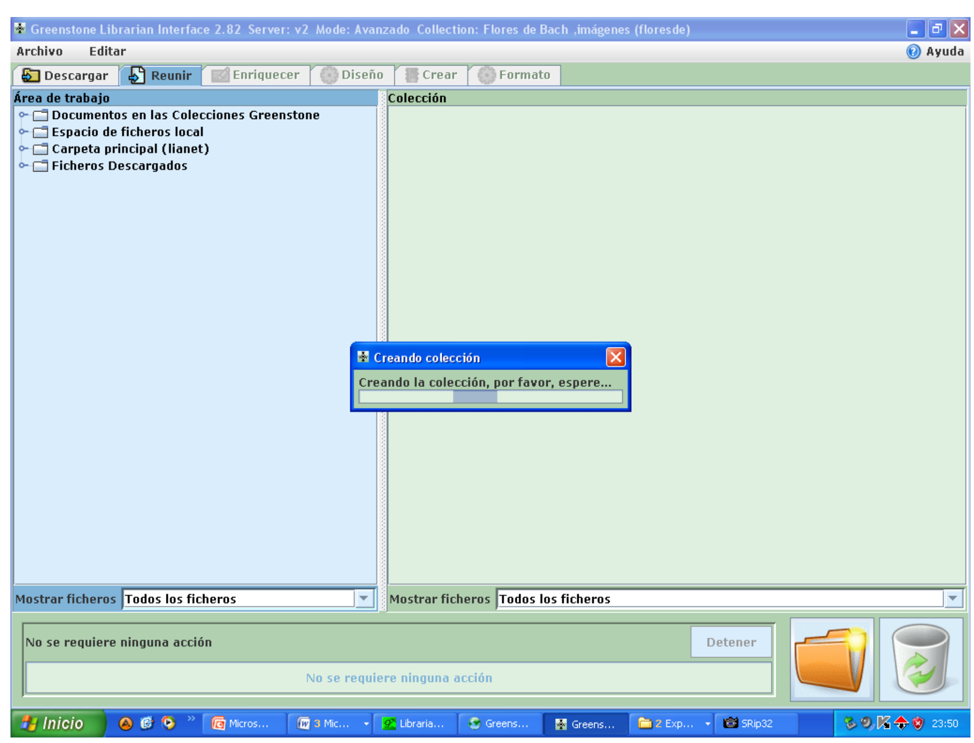


Figura 15.4 Carga de una nueva colección.

Activada la pestaña Reunir localizamos en la computadora la carpeta con los documentos a procesar en este caso la de imágenes; como se observa en la figura el lado izquierdo muestra la carpeta de imágenes, (sombreadas) todos en formato .jpg para la colección. La figura presenta los archivos de imágenes importados en el área de Greenstone. Ver Figura 15.5

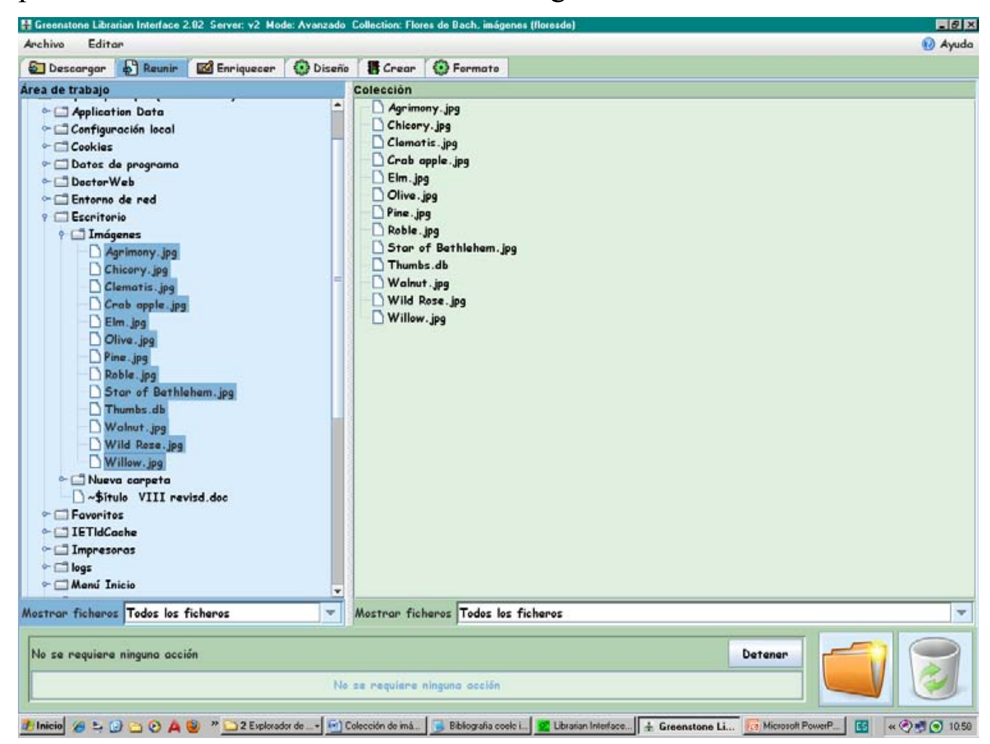


Figura 15.5 Selección de documentos pestaña Reunir.

15.4.1- Pestaña Enriquecer

En la pestaña Enriquecer se le agrega información de los documentos de forma manual para enriquecer los metadatos, se ha seleccionado para este ejercicio, los metadatos Dublin Core: Titulo, Descripción, Formato.

Es importante para esta colección añadir una descripción para cada imagen en forma de metadatos dc. Descripción. Ver Figura 15.6

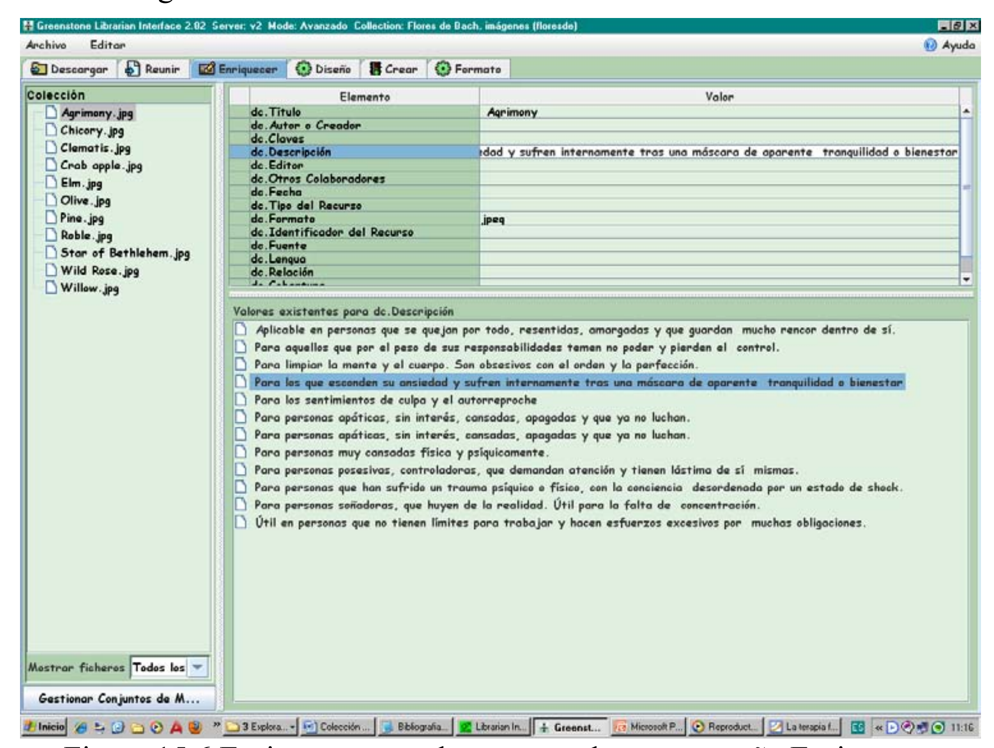


Figura 15.6 Enriquecer metadatos manualmente pestaña Enriquecer.

15.4.2- Pestaña Diseño

El programa ofrece un listado de varios plugins, para este caso como se trabaja con imágenes el más conveniente es *ImagePlugin plugin* pues reconoce todo tipo de formato de imágenes.

En la pestaña Diseño seleccione el *plugin ImagePlugin*, necesitamos acceder a sus “Opciones de configuración” para perfeccionar más la información lo configuramos tal y como se muestra en pantalla la Figura 15.7.

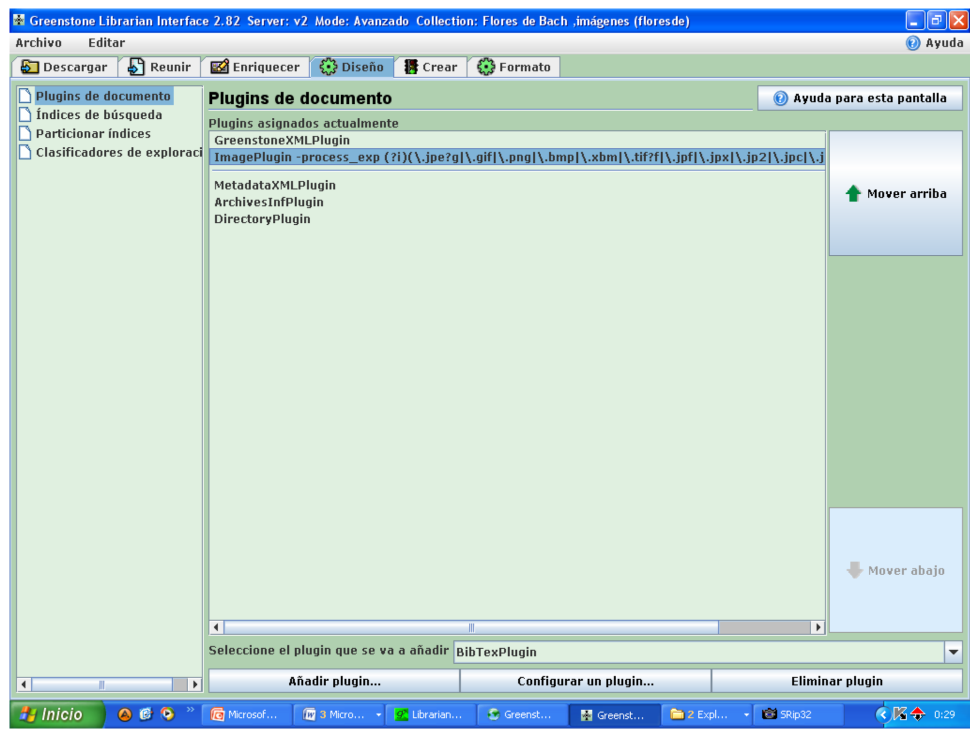


Figura 15.7. Selección y configuración de Plugins.

Otra opción que debe modificarse en la pestaña Diseño es la opción Índices de búsqueda cambiamos los que brinda originalmente el programa y se agregan nuevos; se escogen en la opción Nuevo índices los de Título, Descripción, Ver Figuras 15.7 y 15.8.

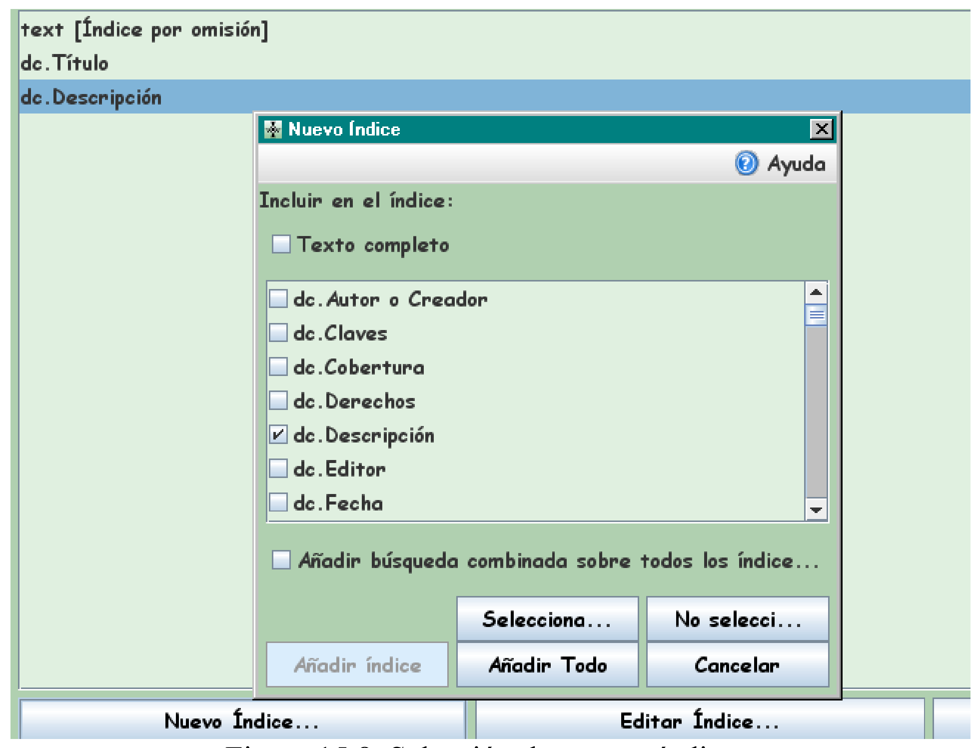


Figura 15.8. Selección de nuevos índices.

En esta colección se trabaja con el indizador MGPP.

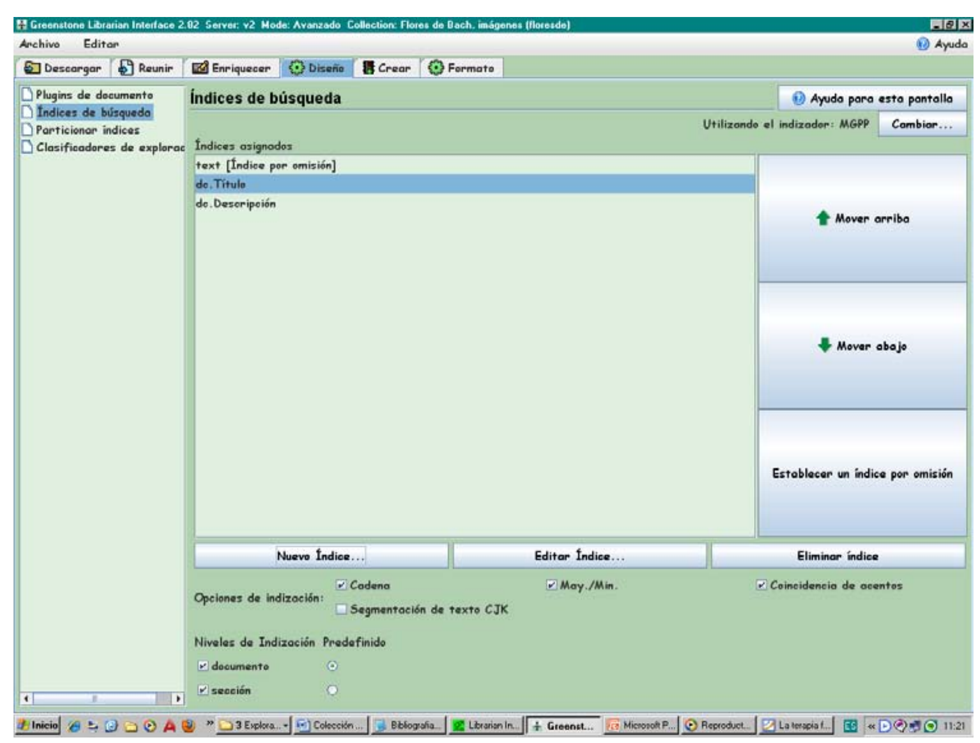


Figura 15.9 Opción índices de Búsqueda.

En cuanto a los clasificadores que se van a configurar para esta colección se modifican los que presenta el software por defecto y se seleccionan Azlist para título y descripción, también seleccione el clasificador *Collage (*Ver Figura 15.9) este último presenta una exhibición de las imágenes en un lienzo de color verde bosque por defecto; pero si desea puede cambiar este color en el argumento bgcolor sustituyendo el código por el que desee; estos cambios se efectúan cuando configura los argumentos del clasificador tal como lo muestra la Figura 15.10.



Figura 15.10. Selección de Clasificadores.

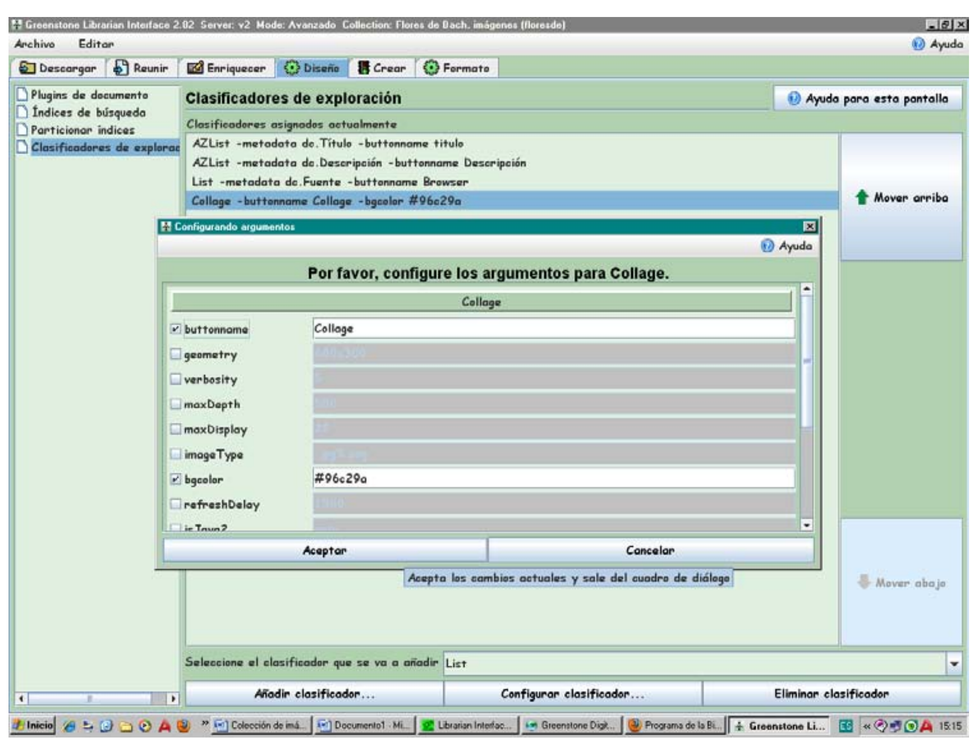


Figura 15.11.Configuración de clasificador Collage.

Este lienzo o refrescador especificado en hexadecimal cambia en milésimas de segundos las diferentes imágenes incluidas en la colección.

Una vez diseñada y configurada la colección se dará paso a la creación de esta, tal como muestra la figura siguiente (Ver Figura 15.12), presenta el mensaje del suceso de importación y procesamiento de los documentos con la pestaña Crear.

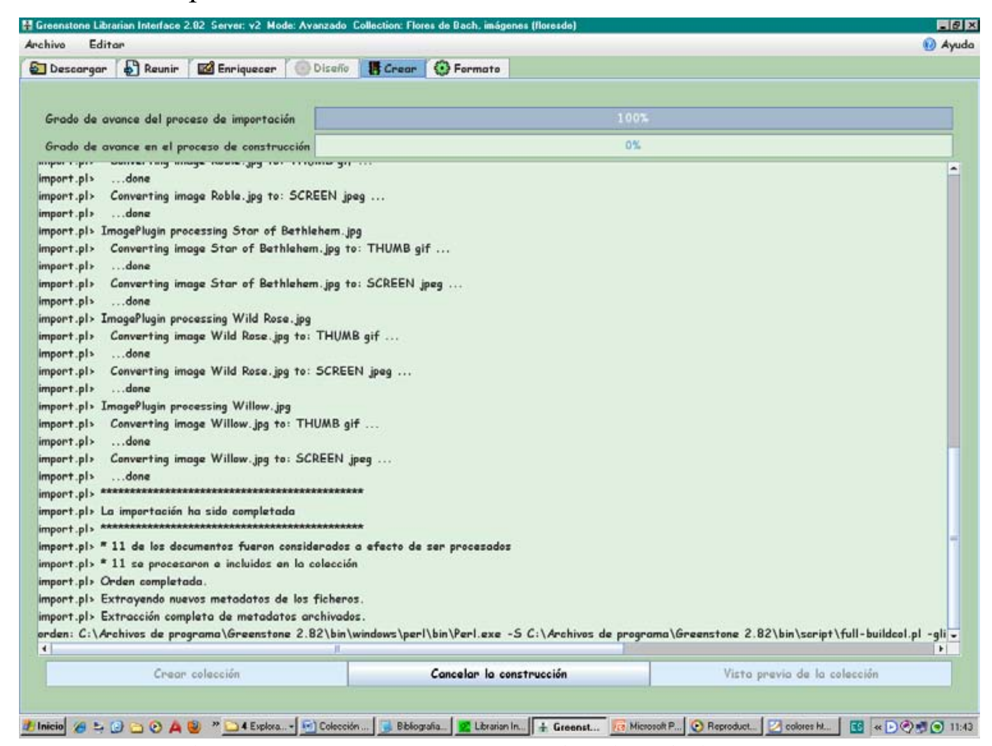


Figura 15.12 Pestaña Crear colección de imágenes.

Esta pestaña le informa al creador de la colección la cantidad de documentos importados y de ellos los que han sido procesados e incluidos en dicha colección; en la figura puede apreciar que los 11 documentos o en este caso las imágenes han sido reconocidos por el plugin seleccionado anteriormente durante el proceso de configuración y por tanto procesados. Ver Figura 15.12.

A continuación aparece en pantalla un cuadro de diálogo. Ver la Figura 15.13, que le indica al usuario que la colección se ha creado y que está lista para verla, con solo dar un clic en la pestaña Vista previa de la colección.

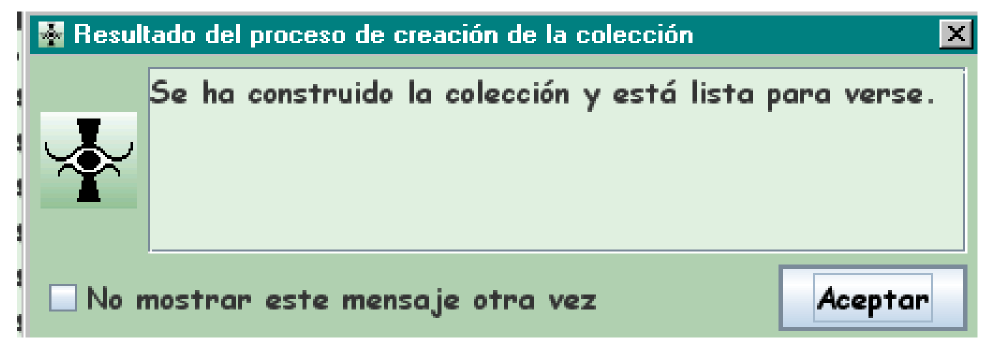


Figura 15.13 Mensaje que indica al usuario que la colección está creada.

A continuación puede revisar cada uno de los botones que aparecen en la barra de navegación para consultar la colección creada.

En pantalla se presenta el botón o Clasificador título, solamente aparece el título que identifica a cada imágenes. Ver Figura 15.14.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0



Figura 15.14. Visualización de Clasificador Título.

Para la creación de esta colección se han trabajado hasta este momento todas las pestañas que contiene GLI, excepto Formato por eso en la figura anterior presenta las imágenes con el nombre original de los documentos y el icono de HTML, no útiles para este tipo de colección.

Para lograr la vista deseada, hemos de modificar las cadenas de formatos presentes en la opción Formato dentro de la pestaña mencionada anteriormente, por eso para el Clasificador de título (CL1) optamos por la siguiente cadena de formato (4):

```
<td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
<b><u>{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}</b></u>
[/highlight]
<br>
{Or}{[dc.Description]<b>dc.Descripción</b>}<td valign="left">[highlight] [srclink]
</td>
```

El resultado será como lo presenta la siguiente Figura 15.15.



Figura 15.15. Colección de imágenes en la que se ha añadido el Clasificador de título (CL1).

A continuación les mostraremos en la Figura 15.14 con el clasificador Descripción como es posible visualizar un comentarla de cada una de las imágenes y se describen las características de cada flor escogida, de esta forma le permite al usuario tener información muy precisa del uso de cada flor para el tratamiento del estado anímico o la forma de ser de acuerdo con las necesidades de cada persona. Para lograr este efecto se debe incluir en la cadena de formato el metadato dc.Descripcion, por eso hemos seleccionado la siguiente cadena de formato (5):

```
<td valign="top">[ex.srclink]{Or}{[ex.thumbicon],[ex.srcicon]}[ex./srclink]</td>
<td valign="top">[highlight]
<b><u>{Or}{[dc.Title],[exp.Title],[ex.Title],Untitled}</b></u>
```

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

```
[/highlight]
<br>
{Or}{[dc.Description]<b>dc.Descripción</b>}<td valign="left">[highlight] [srclink]
</td>
```

El resultado será como lo presenta la Figura 15.14



Figura 15.16 .Clasificador Descripción en la colección de imágenes.

Con solo dar un clic en este clasificador se presenta una pantalla que muestra este lienzo que se visualiza en la figura de color verde bosque en forma de collage todas las imágenes procesadas y las presenta de forma rápida en cuestiones de milésimas de segundos. Ver Figura 15.16.



Figura 15.17 Visualización de Clasificador Collage.

Si en cambio desea reducir el tamaño de las imágenes puede configurar en el plugin seleccionado tamaño imagencita, ajústelo a 50, y haga clic en <OK>.Para lograr el cambio que surte este efecto debe de nuevo crear la colección en la pestaña Crear.

Como consecuencia de este cambio el efecto resultante es el que muestra la siguiente Figura 15.18 con las imágenes en tamaño más pequeño.

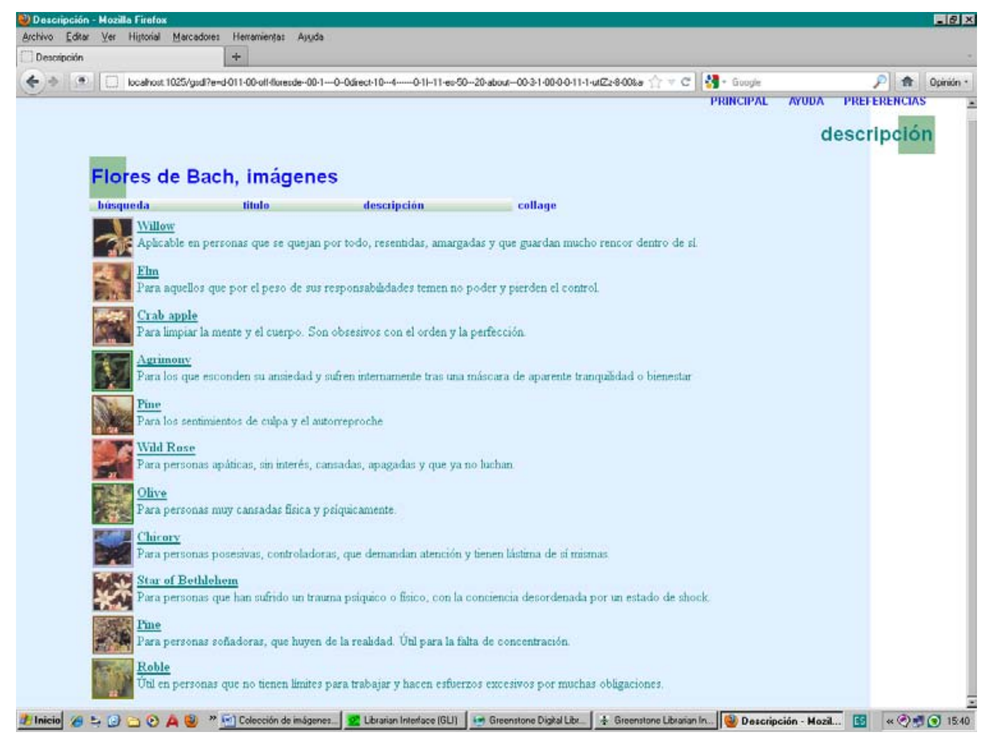


Figura 15.18. Configuración del plugin seleccionado con ajuste de imagen.

15.5- Consideraciones Finales

El ejercicio puntualizado en este capítulo ha demostrado que este software con las diferentes herramientas que posee; promete opciones diversas para crear colecciones digitales atractivas y adaptarlas a las necesidades de usuarios.

15.6- Referencias Bibliográficas

(1) IFLA. Directrices para proyectos de digitalización. [Sitio Web en Internet] 2002 . Disponible en <http://archive.ifla.org/VII/s19/pubs/digit-guide-es.pdf>

(2) Biblioteca digital Greenstone. Guía del programador Capitulo 2 Explotación óptima de sus documentos. [Sitio Web en Internet] Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_getting_the_most_out_of_your_documents.htm>

(3) Biblioteca digital Greenstone. Guía del usuario Capitulo 3. Creación de colecciones Greenstone. [Sitio Web en Internet] Disponible en <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_making_greenstone_collections.htm>

(4) Bainbridge D, McKay D, Witten H. I. Guía del programador. Biblioteca digital Greenstone.

[Sitio Web en Internet].Departamento de Informática. Universidad de Waikato, Nueva Zelanda. Marzo de 2004.Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Develop_es_all.html>

(5) Wiki.Greenstone . Greenstone tutorial exercise. A simple image collection. [Website in Internet]. 2010. Available from::<http://greenstone.sourceforge.net/wiki/index.php/Tutorial_exercises>

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

CAPÍTULO 16

16- Creación de una colección multimedia

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH)

Palabras clave: Bibliotecas Digitales, Desarrollo Sostenible, Economía Digital, Estrategias de Inclusión Digital. , Sociedad de la Información, Software Libre, UNESCO.

16.1- Introducción

Las colecciones Greenstone pueden contener texto, imágenes, sonidos y videoclips. La mayor parte del material no textual está enlazado a los documentos textuales o va acompañado de descripciones textuales (por ejemplo, las leyendas de las figuras) para permitir la búsqueda y consulta en texto completo.

16.2- Establecer la nueva colección. Seleccionar los documentos a incluir

Comenzaremos una nueva colección de multimedia a la que llamaremos "small beattles "

Como estamos ya habituados nos dirigiremos al menú file y allí indicaremos new para comenzar a crear una nueva colección (1). Ver Figura 16.1.

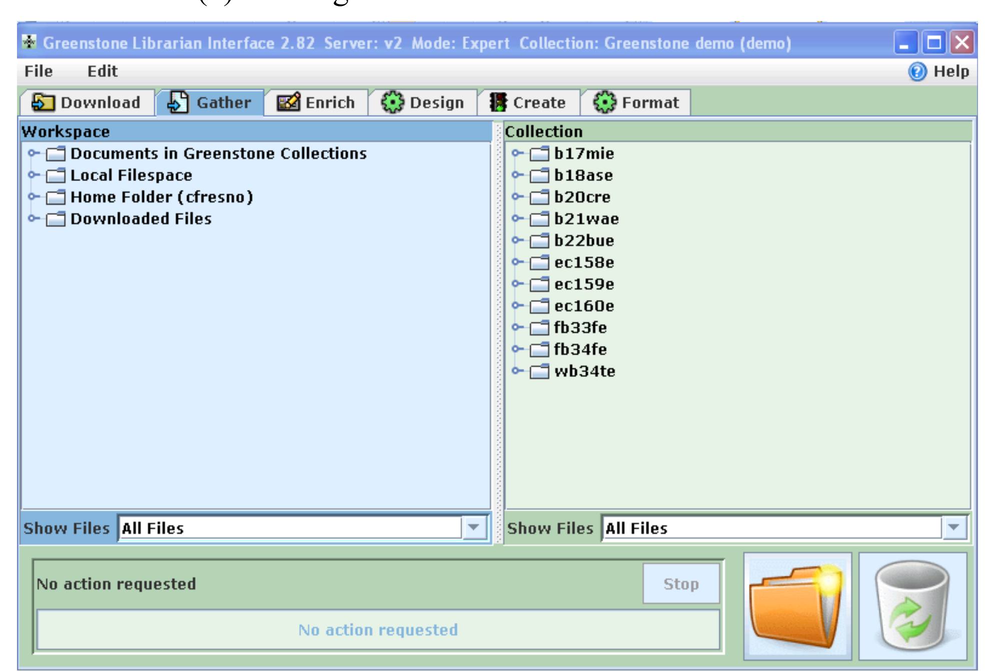


Figura 16.1. Inclusión de los documentos a incluir en la colección small beatles.

En la Figura 16.2, se destaca la pantalla donde referiremos los datos que calificaran la colección que crearemos.



Figura 16.2. Pantalla donde identificaremos las características de la nueva colección.

En la Figura 16.3, identificaremos el contenido de la carpeta con la que vamos a trabajar que está ubicada en el disco D de la computadora.

Figura 16.3. Contenido de la carpeta con los documentos para la colección *small Beatles*.

En el proceso de traslado de los materiales que se van a incluir en la colección hacia el panel verde, destacaremos que aparece un cuadro de dialogo que nos pregunta si deseamos añadir un pluggin para mp3 y otro para MARC, pues son requerimientos que automáticamente requiere el programa al detectar la presencia de documentos de este tipo.

La Biblioteca del Congreso ha desarrollado MARC Standards Office que permiye establecer una red y utilizar los datos en MARC en un ambiente HTML. Esta red resulta flexible y extensible a una gran cantidad de usuarios en el mundo que trabajan con datos MARC en la medida de sus necesidades (2). Esta red puede contener varios componentes tales como esquemas, hojas de estilo, y herramientas de software desarrolladas y mantenidas por la Biblioteca del Congreso. <http://www.loc.gov/marc/> MARC.

16.3- Selección de los metadatos y creación de la colección

Nos dirigiremos al panel enriquecer para agregar valores a los metadatos. Ver Figura 16.4.

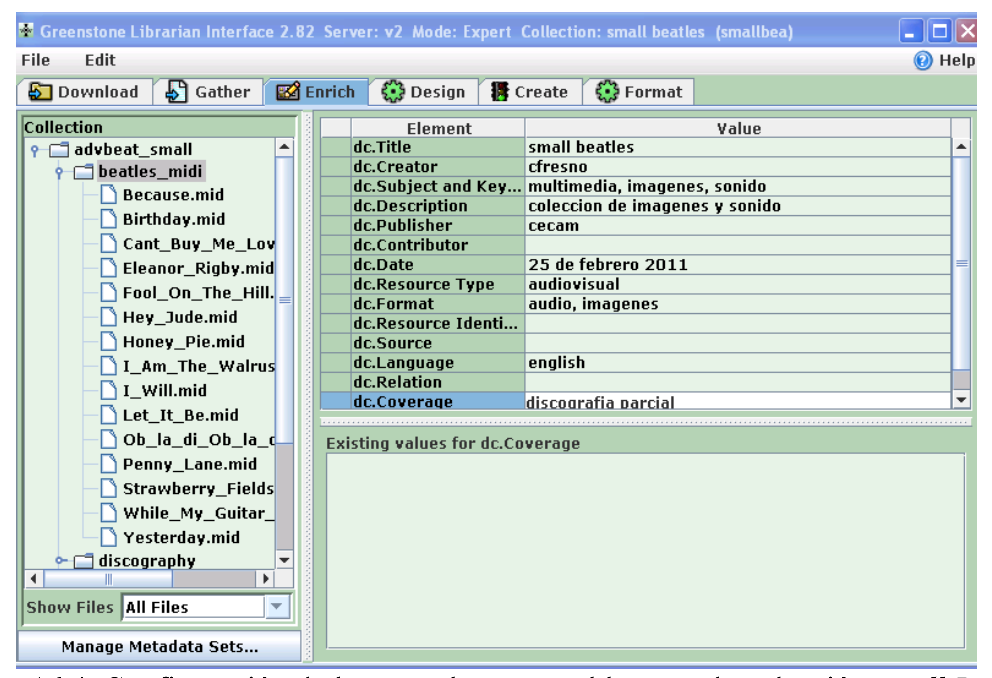


Figura 16.4. Configuración de los metadatos a establecer en la colección *small Beatles*.

Pasar al botón crear para dar inicio a la creación de la colección. Como es habitual podrán seguir el curso del progreso de la construcción de la colección, hasta que aparezca el cuadro de dialogo que indique que la colección ha sido creada satisfactoriamente. Ver Figura 16.5.

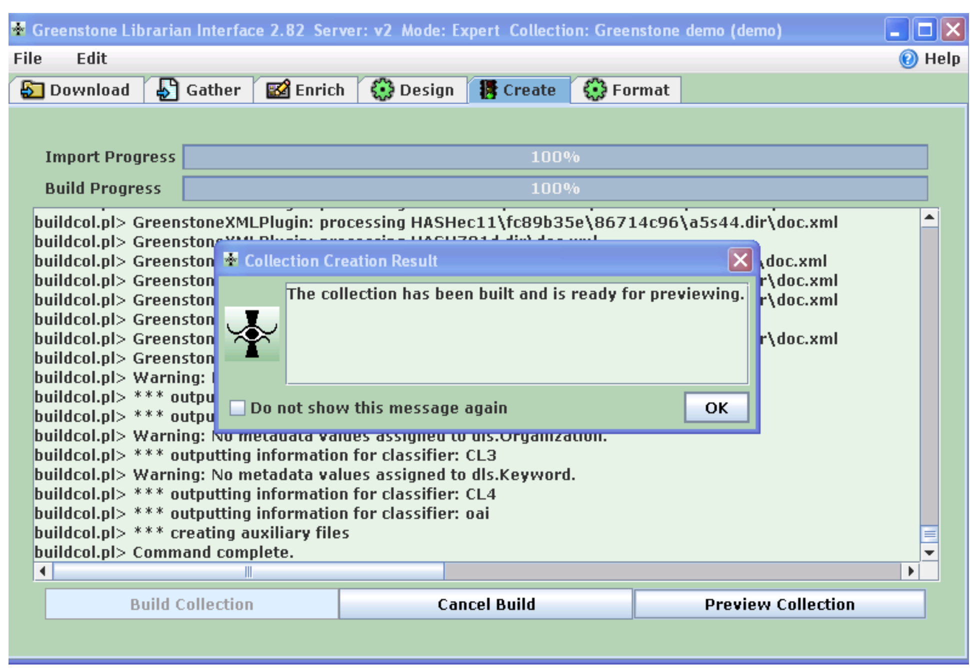


Figura 16.5. Proceso de creación de la colección *small Beatles*.

Nos dirigiremos al botón file y allí indicaremos Open. De ese modo podremos visualizar todas las colecciones que contiene el programa incluyendo por supuesto la recientemente creada por nosotros. Ver Figura 16.6.

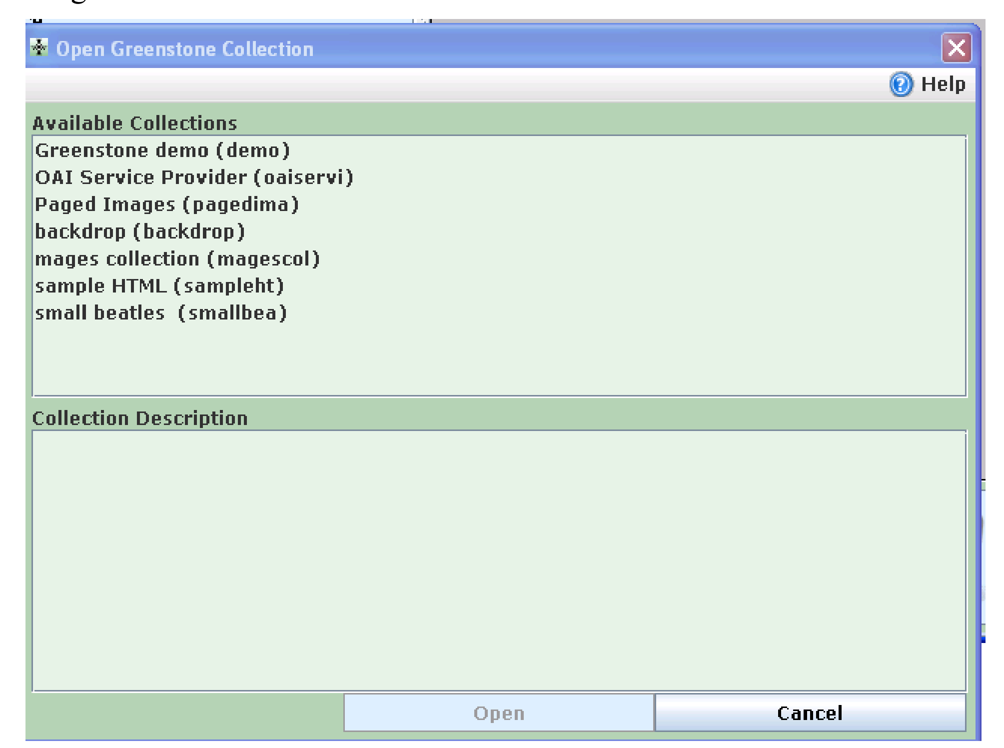


Figura 16.6.Identificación de la colección *small Beatles* entre las colecciones disponibles.

Si nos introducimos ahora en la interfaz de usuario podremos visualizar la colección de multimedia que hemos creado, Queremos destacar que en esta colección hasta el momento solo hemos seleccionado como clasificadores para orientar la búsqueda *titles y filenames*. Ver Figura 16.7

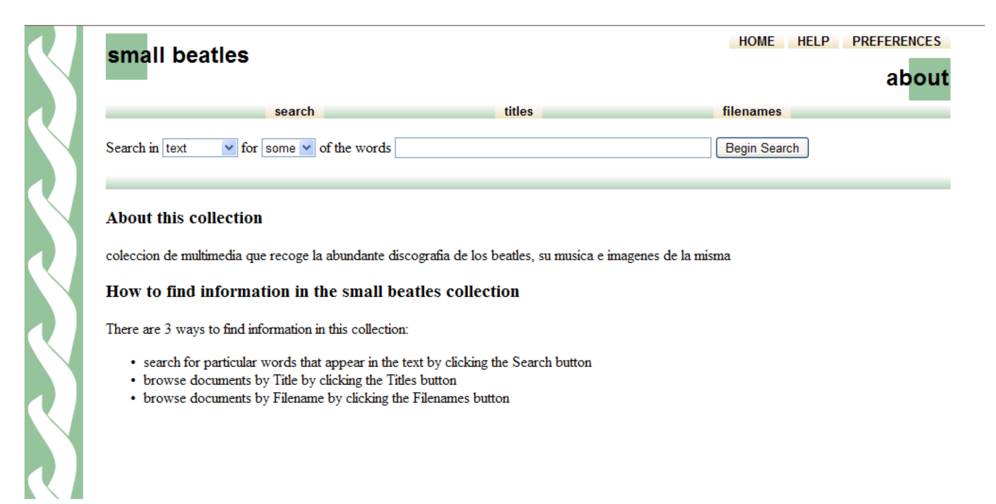


Figura 16.7. Interfaz de usuario y presentación de la colección *small Beatles*.

Seleccionamos el clasificador titile por ejemplo e introducimos la palabra " yesterday ". Seleccionamos la opción all. Y este será el resultado que obtendremos. Destaquen que entre los archivos recuperados encontramos archivos mp3, htm y marc . Ver Figura 16.8.

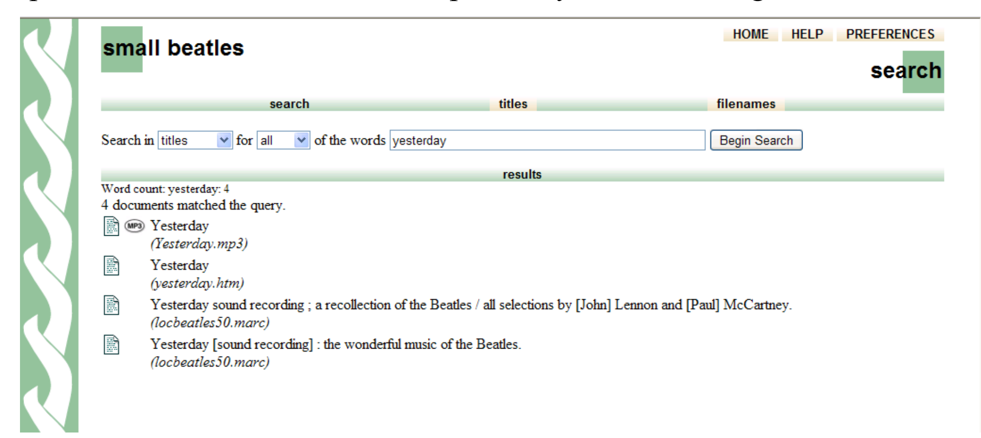


Figura 16.8. Resultado de la búsqueda en la colección *small Beatles* por el clasificador "*title*"

16.4- Cambiar aspectos del formato de la colección. Uso del Clasificador PHIND

Ahora vamos a incorporar el Clasificador PHIND en esta colección. Nos dirigiremos al panel Diseño a la sección de Browsing Clasifiers y en el botón *Add Clasifier agregar Phind Classifier.* Ver Figura 16.9.

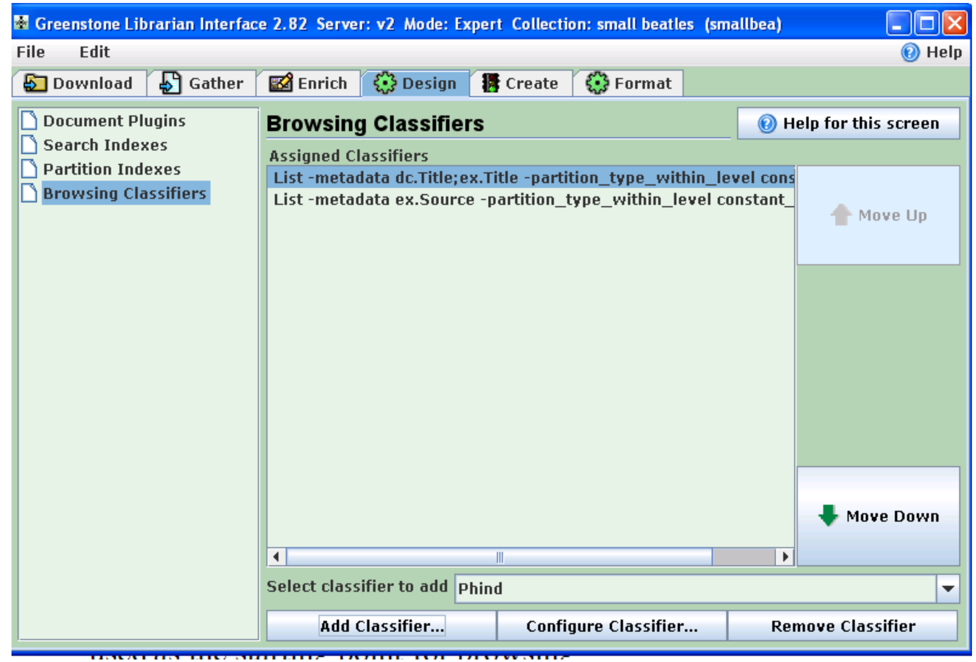


Figura 16.9.Seleccion del clasificador "Phind "para incluir en la colección *small Beatles*.

Al agregar el clasificador Phind aparecerá un cuadro de dialogo como este en el que Ud. debe configurar los argumentos necesarios. Para ello marque el cuadriculado text y aparecerá section title. Agregue allí section text. Ver Figura 16.10.

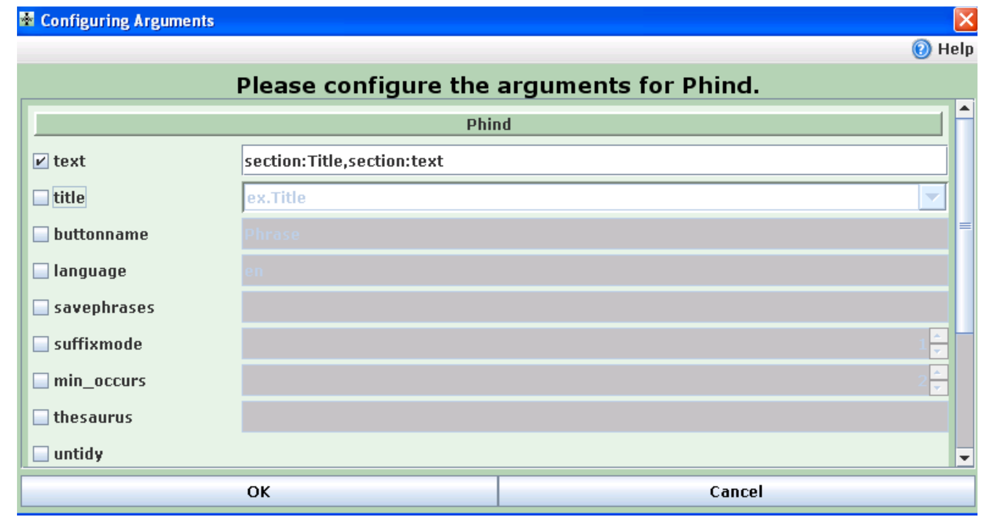


Figura 16.10.Configuracion de los argumentos necesarios para el clasificador PHIND en la colección *small Beatles*.

Al regresar al panel Diseño podremos observar como aparece agregado en nuestra colección el clasificador Phind. Ver Figura 16.11

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

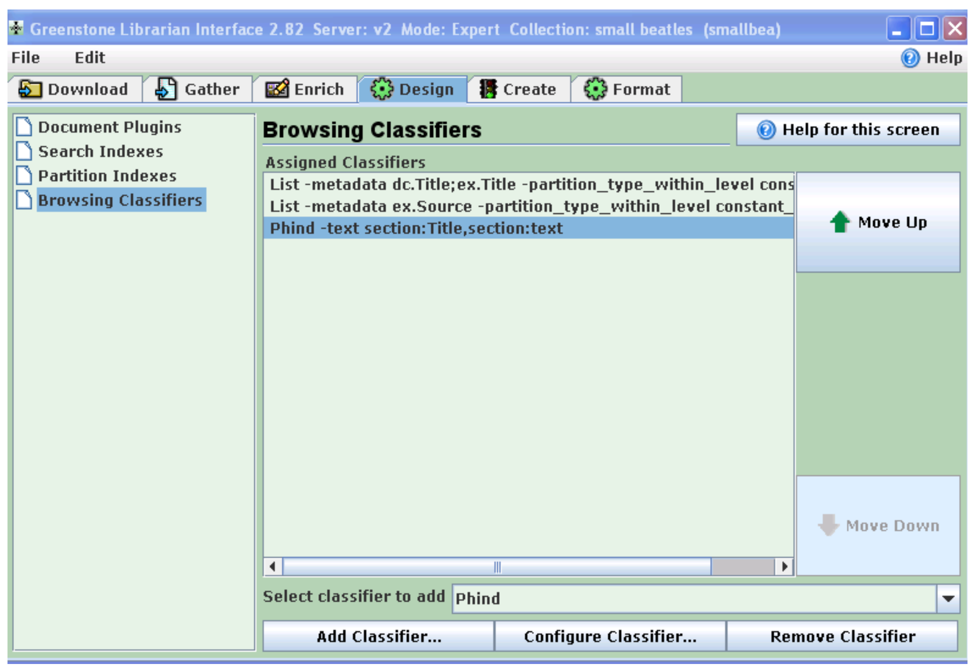


Figura 16.11. Visualización del clasificador PHIND entre los clasificadores seleccionados en la colección *small Beatles*.

Veamos ahora los cambios que se producen en la interfaz de usuario. Aquí podemos observar como aparece el nuevo clasificador "Phrases". Este clasificador si bien funciona mejor con palabras simples resulta de mucha utilidad para la búsqueda en grandes colecciones (3). Ver Figura 16.12.

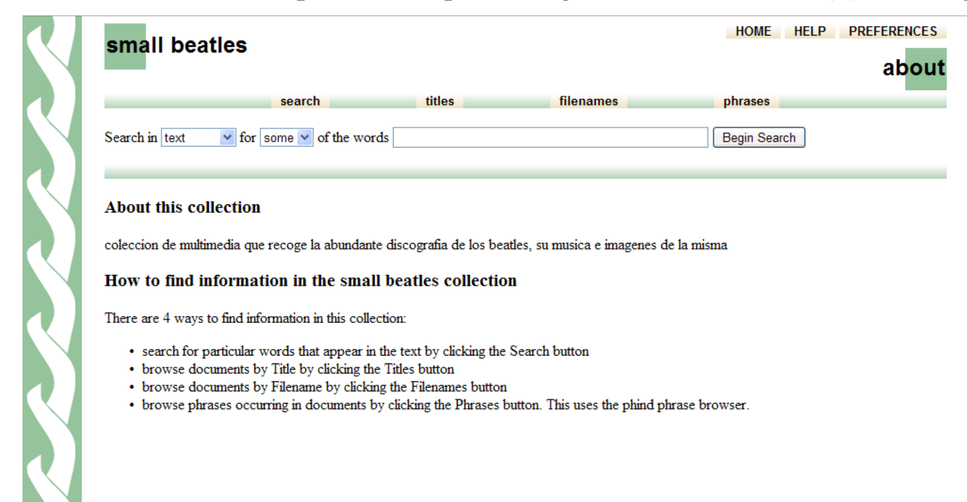


Figura 16.12. Visualización del nuevo clasificador *Phrases.*

Ahora por ejemplo procedamos a iniciar la búsqueda con la palabra "band" y comprobar los resultados por ustedes mismos.

16.5- Insertar una figura que identifique la colección

Nos dirigiremos al panel formato y en la sección general, y situándonos en la opción URL to 'about page' image, con el botón browse, buscaremos y seleccionaremos la siguiente imagen que encontraremos en la computadora, siguiendo este camino, como ejemplo, pero puedn dirigirse a cualquier sitio en u computadora y seleccional la imagen que deseen: **sample_files** → **beatles** → **advbeat_lsamall** →**images** →**The –Real –Penny-Lane**. Ver Figura 16.13.

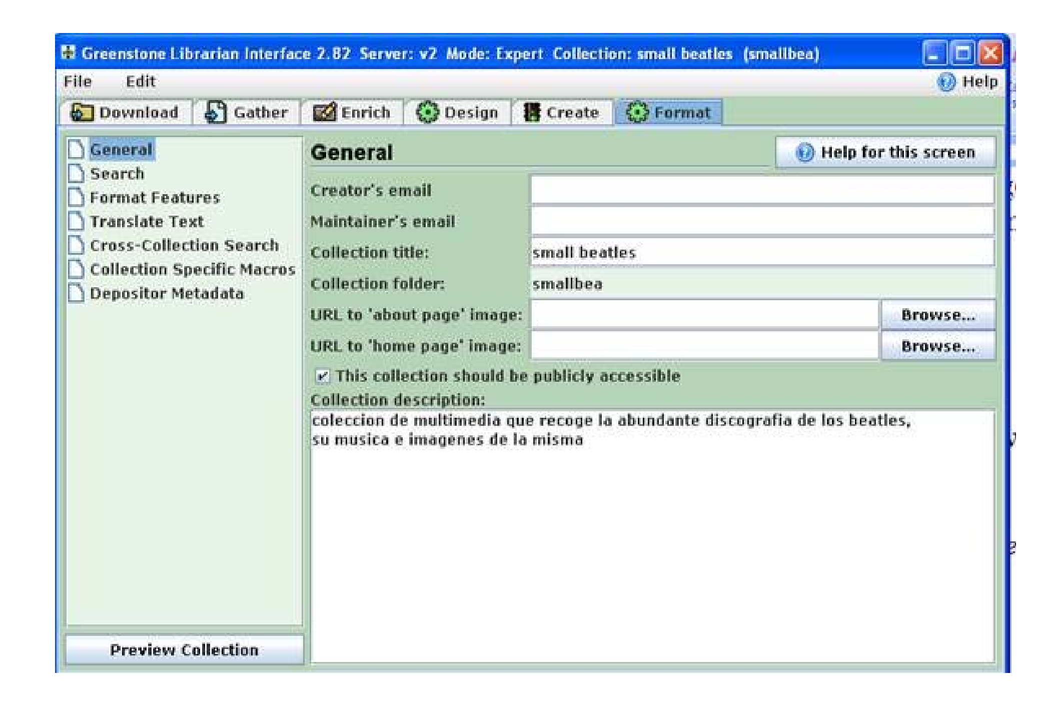


Figura 16.13. Opción selección de una imagen que identifique la colección
Figura 16.14. Inserción de una figura que caracteriza la colección "*small beatles*"



Y podremos contar con la imagen que prefieran para caracterizar la colección. Por supuesto pueden recodar la clase con Snagit que dieron en el curso básico y modificar la dimensión de la imagen.

Nos dirigiremos ahora al panel Diseño y allí buscaremos en select pluggin to add el plugin deseado y lo añadiremos utilizando Add Plugin

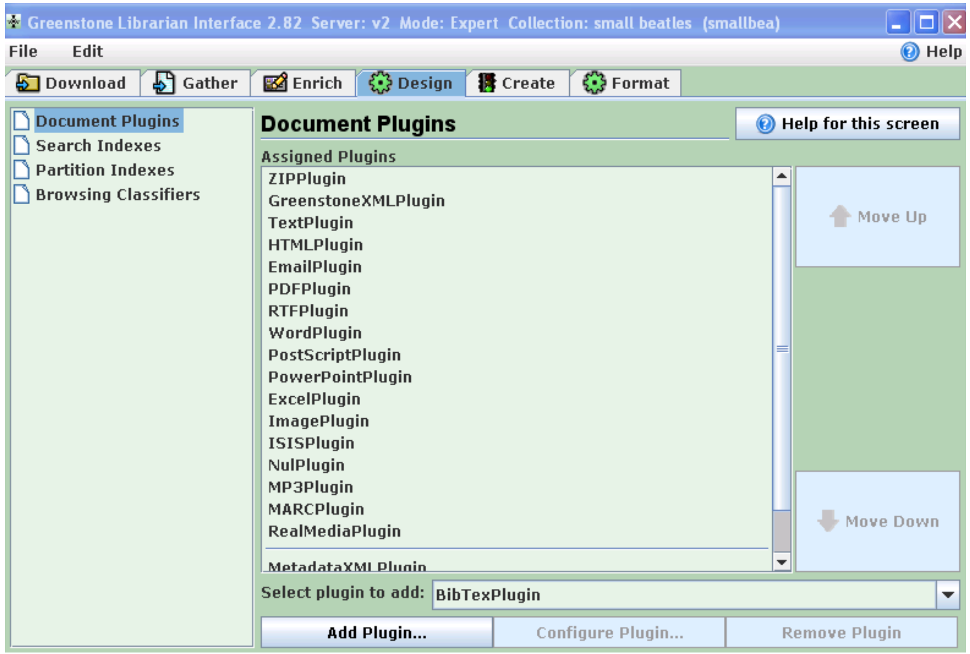


Figura 16.15. Adicionar un nuevo plugin a la colección “small beatles”.

Una vez presionado el botón Add Plugin aparecerá este cuadro de dialogo.

Nos dirigiremos al panel Diseño y allí buscaremos en "select pluggin to add" el plugin deseado y lo añadiremos utilizando Add Plugin. Una vez presionado el botón Add Plugin aparecerá este cuadro de dialogo. Ver Figura 16.15.

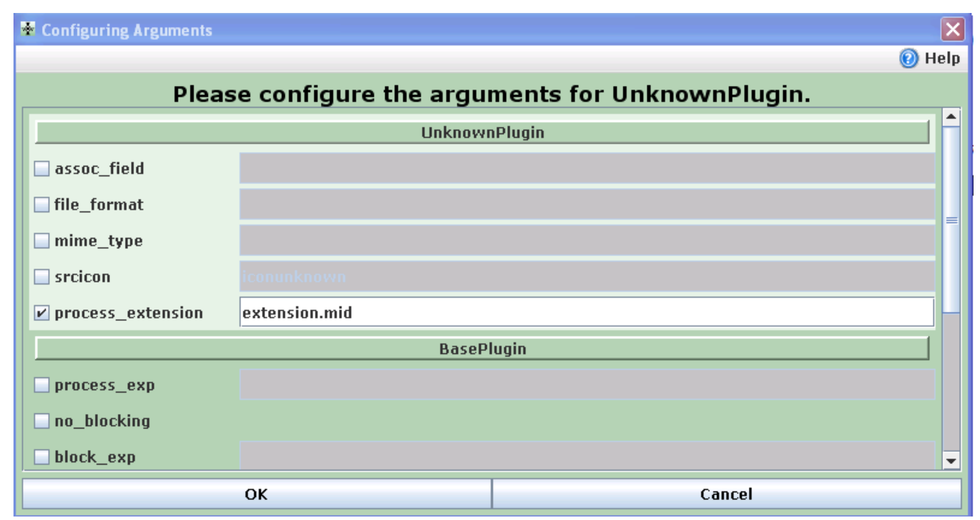


Figura 16.16. Opción selección de una imagen que identifique la colección.

Active las casillas **file-format** a **process extension** y asigne los datos que se indican a continuación. Ver Fig. 16.17

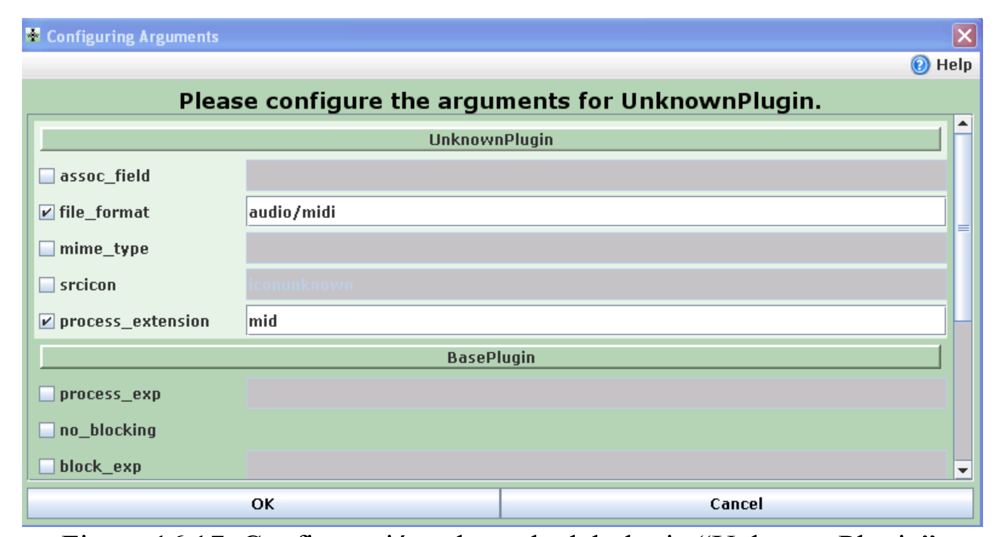


Figura 16.17. Configuración adecuada del plugin “UnknownPlugin”.

Cuando estemos configurando los argumentos para este plugin activaremos la casilla process-extensión y allí escribiremos mid para que este plugin nos ayude a procesar los formatos midi, Activaremos también la casilla File-Format y escribremos allí audio/midi. Cree de nuevo la colección y vera como puede reconocer ahora los archivos midi. Ver Figura 16.17.

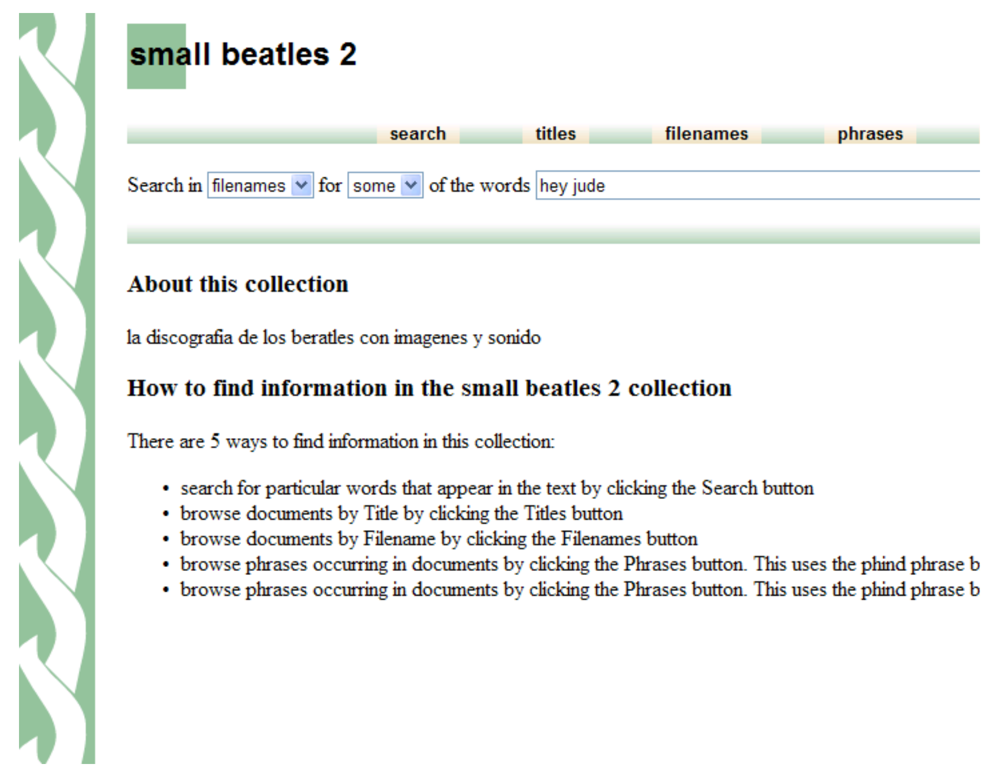


Figura 16.18. Configuración de los argumentos del UnknownPlugin.

Visualización de la interfaz de usuario después de añadir el UnknownPlugin y añadir Búsqueda en la colección "*small beatles 2*". Ver Figura 16.18

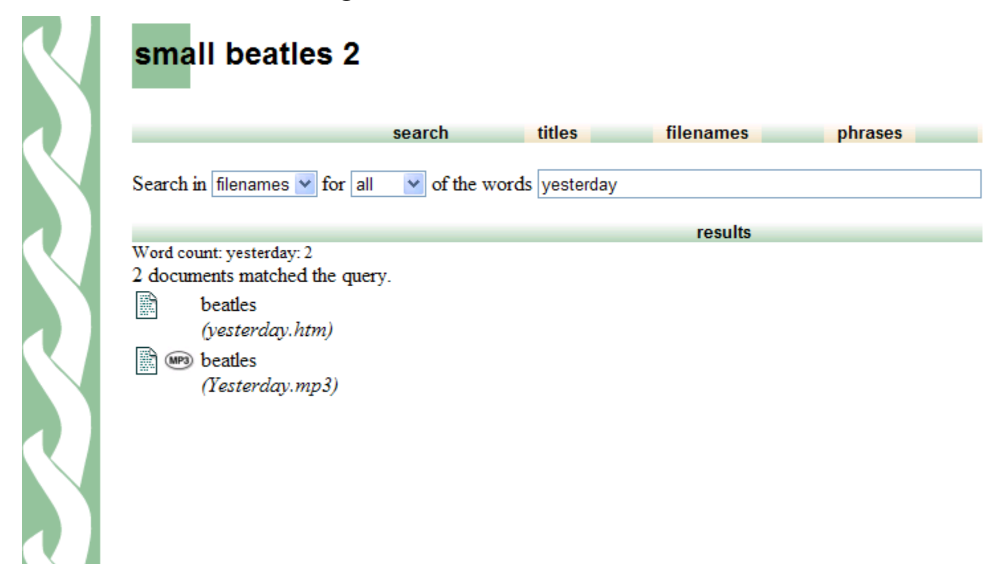


Figura 16.19. Visualización de la colección "small beatles" ahora denominada "*small beatles 2*", tras haber añadido el "*UnknownPlugin*".

Ud. Podrá consultar un pequeño instructivo básico de cómo proceder en la creación de una colección multimedia si se dirige a la sección Anexos de este libro.

16.6- Consideraciones Finales

Como han podido apreciar no resulta nada difícil crear una colección multimedia básica y la opción de mejorar el rendimiento en la recuperación de documentos al interior de la misma dependerá de

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

las opciones que UD. elija.

16.7- Referencias Bibliográficas

(1) wiki.greenstone.org .Greenstone tutorial exercise. Building a multimedia collection [Website in Internet]. 2010.Available from:<http://wiki.greenstone.org/wiki/gsdoc/tutorial/en/multimedia_collection.htm >

(2) Méndez E. Senso A. Introducción a los metadatos. Estándares y aplicación Conceptos asociados al uso de meta información . Unidad de Autoformación.[Sitio Web en Internet] SEDIC 2004. Disponible en: <http:www.sedic.es >

(3) Greenstone.org Biblioteca Digital Greenstone; Guía del Programador. Capitulo 2 Explotación óptima de sus documentos. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/style.css >

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

CAPÍTULO 17

17- Publicación de las colecciones: exportación de colecciones mediante las opciones Exportar y Exportar a CD ROM desde el menú Archivo

Lic. Elvia Vázquez Velázquez (elvia@cecam.sld.cu) Licenciada en Letras.
Centro de Cibernética Aplicada a la Medicina (CECAM) I.S.C.M.H. Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354

Palabras clave: Bibliotecas Digitales, Colecciones de documentos, Greenstone, Formato DSpace, Exportar colecciones.

17.1- Introducción

El objetivo principal que se busca al implementar una biblioteca digital es maximizar la visibilidad, el uso e impacto de la producción científica y académica de una institución y sus investigadores; además de facilitar, mantener y preservar el acceso a la información. (1)

Primordial resulta el hecho de promover y mantener el acervo de información incluida en colecciones tanto en bibliotecas o repositorios institucionales y preservar esta información como medida de seguridad y conservación. Una de las formas posibles de asegurar esto acuerda valerse de programas que permitan exportar las colecciones digitales a diferentes formatos estándares y a CD-ROM/DVD.

Uno de ellos es Greenstone, que puede exportar los contenidos y/o metadatos de una colección a varios formatos estándares, incluyendo METS, DSpace y MARC-XML (2) así como también a CD-ROM/DVD.

17.2- Características Dspace

Posee mecanismos para validar y controlar el acceso a: Metadatos, documentos, es interoperable, implementa el protocolo Open Archives *Initiative Protocol for Metadata Harvesting*, responde a las OpenURLs, almacena cualquier tipo de documento digital en cualquier formato (4).

Existen opciones específicas para los diferentes formatos. Indique que ficheros XSLT deben aplicarse a los documentos XML resultantes para personalizar el formato de salida. Para exportar a MARCXML se usa un fichero de mapeado para pasar los metadatos de Greenstone a campos MARC. El fichero de mapeo predefinido sólo exporta a Dublin Core. Puede indicar un fichero propio a utilizar. (3).

Según informa el Grupo de la Red Latinoamericana de Greenstone en la versión 2.82 el plugin MARC (XML) ahora importa como Dublin Core, antes que como metadatos extraídos. Los archivos de mapeo marc2 (q)dc.txt han sido modificados para manejar esto. Las reformulaciones de viejas colecciones MARC necesitarán actualizar sus declaraciones de índice/clasificador/formato para trabajar con estos nuevos metadatos (5)

Esto es posible hacerlo desde el menú de la Interfaz del Bibliotecario, que tiene una opción llamada Exportar que permite exportar colecciones en diferentes formas. Ver Figura 17.1

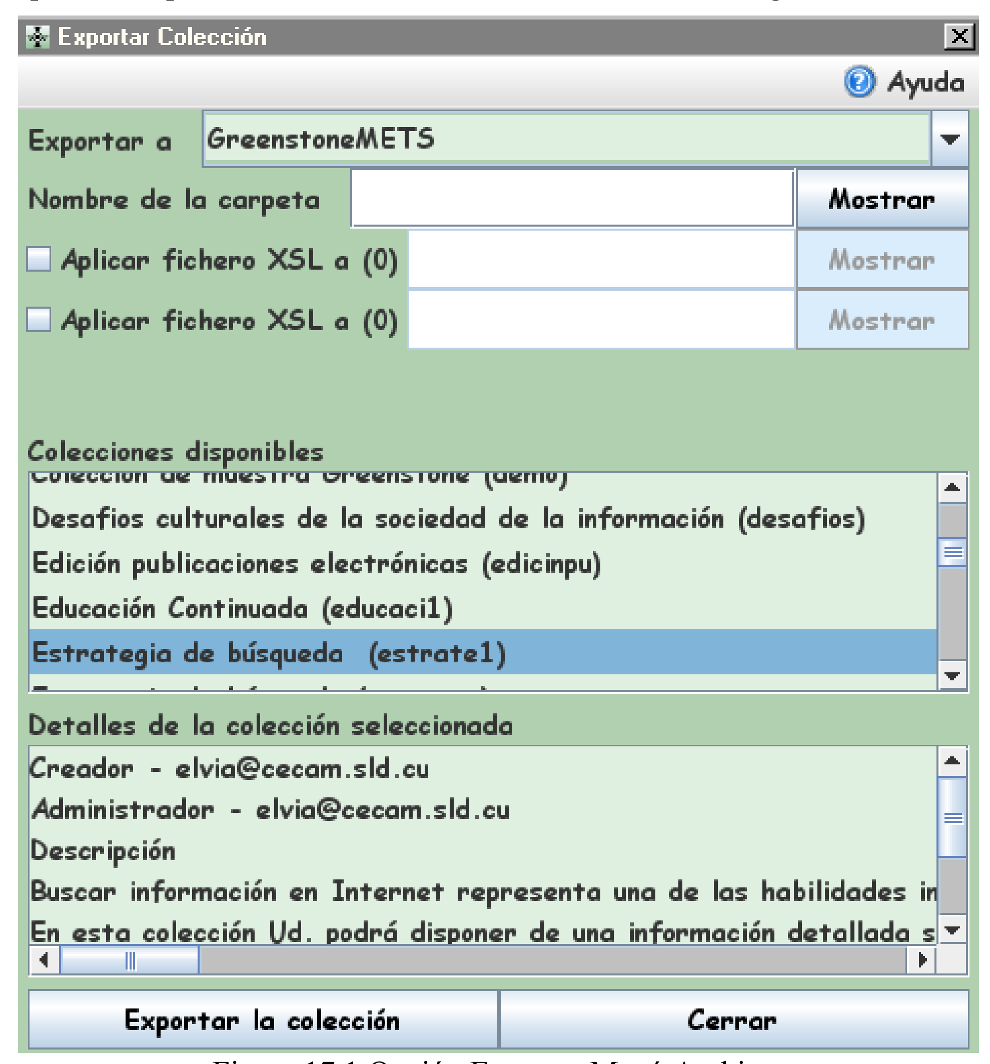


Figura 17.1 Opción Exportar Menú Archivo.

Para exportar una o más colecciones abra el menú Archivo y elija la opción "Exportar...". Puede elegir en qué formato va a exportar de la lista desplegable, debe indicar un nombre para el directorio en el que quiere ubicar los ficheros exportados; los ficheros terminarán en **<ruta a greenstone … xxx>**, este **xxx** es el nombre que Ud. indicó. Seleccione una o varias colecciones de la lista de colecciones disponibles en su biblioteca, y pulse sobre "Exportar colección (es)".

La exportación en Greenstone es un proceso aditivo, los nuevos archivos exportados estarían agregados con diferentes nombres de carpeta a aquellos que ya están en la carpeta de exportación (6).

17.3- Publicación de las colecciones en CD-ROM/DVD

Una vez finalizado el proceso de creación de una colección se puede obtener una copia o salva de la colección en un CD o DVD; para ello abra el menú “Archivo” y seleccione la opción Exportar a un CD-ROM/DVD. Ver Figura 17.2

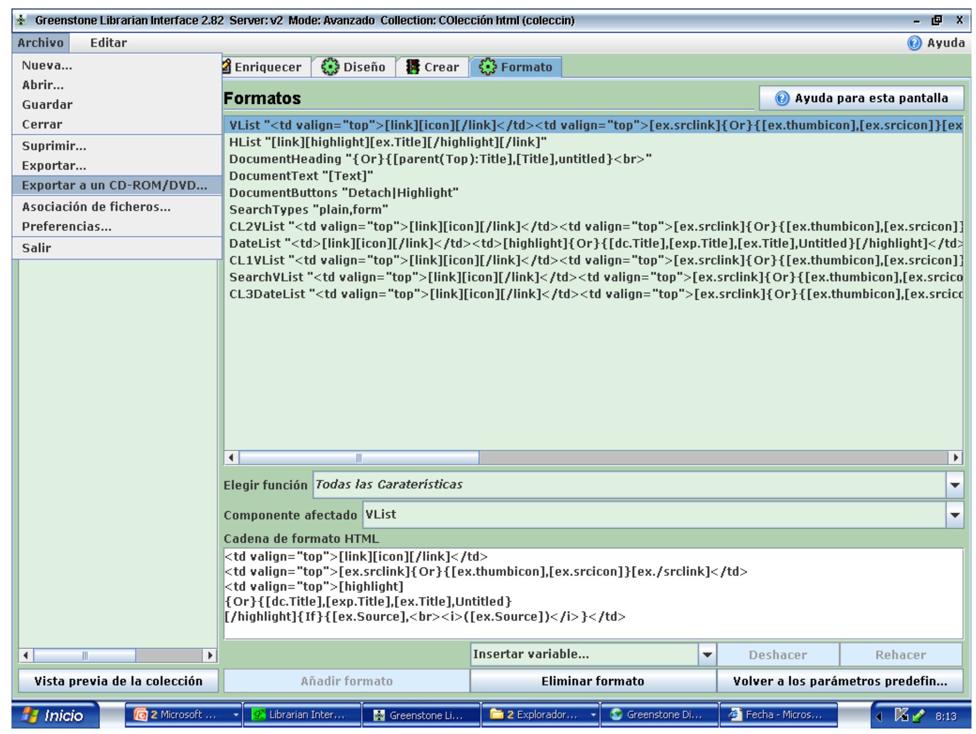


Figura 17.2 Menú Archivo opción Exportar a CD-ROM/DVD.

Al instante aparece una lista de las colecciones creadas y usted selecciona la que desee exportar; para el ejercicio hemos elegido la colección de imágenes creada anteriormente en el capítulo 16, al seleccionarla o marcarla con un shake mark verá su descripción en la parte inferior en la opción Detalles de la colección, debe optar por una de las dos opciones que le indica este recuadro ya sea si instalará algunos ficheros cuando se use y que resulta más rápido o si no instalará ningún fichero; posteriormente debe designar un nombre para el disco y a continuación haga clic en “Exportar”. Ver Figura 17.3

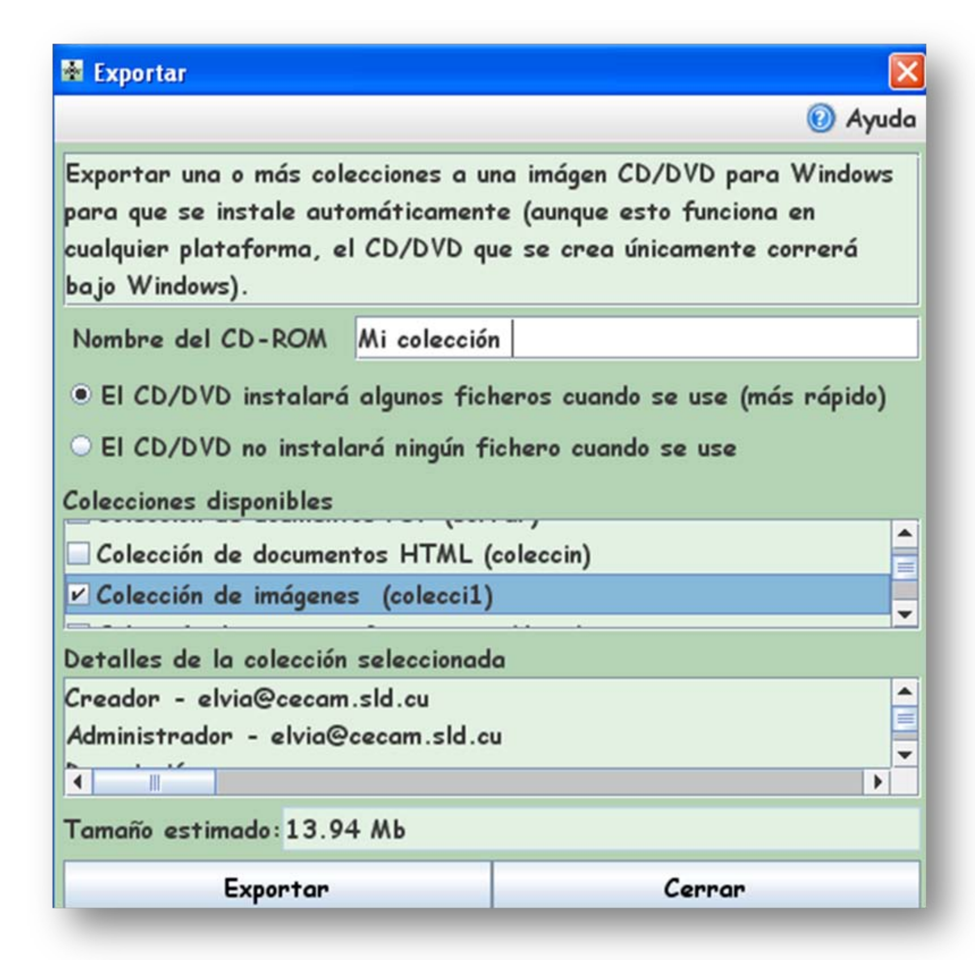


Figura 17.3 Opciones para seleccionar la colección a exportar.

Se produce la copia de varios archivos, los incluidos en la colección, para lo cual se necesitan de algunos minutos, surge un pequeño cuadro de diálogo que se lo informa tal cómo se puede ver en la Figura 17.4



Figura 17.4 Mensaje que informa sobre la copia de los archivos que incluyen la colección.

La pantalla muestra el resultado final de la exportación de la colección que le informa que revise los resultados de esa exportación y le muestra cuál es la ruta para ello; **C:\Archivos de programa\Greenstone2.82\tmp\exported_collections**.

Si usted no especificó un nombre para el CD-ROM, entonces el nombre del archivo será **exported_collections** en vez de **exported_imagencollections**. (Ver Figura 5)

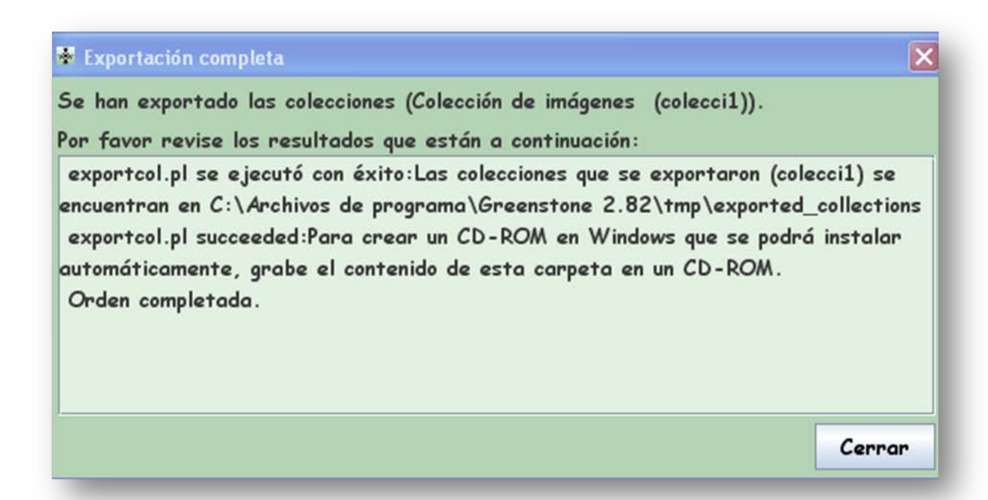


Figura 17.5 Mensaje que indica que la colección ha sido exportada.

Cuando usted revisa la ruta indicada en su computadora la pantalla muestra el resultado final de la exportación de la colección que se encuentra dentro de la carpeta **tmp** dentro de la estructura interna de Greenstone. La carpeta que genera la exportación se nombra: **exported_collections** tal como lo muestra la Figura 17.6

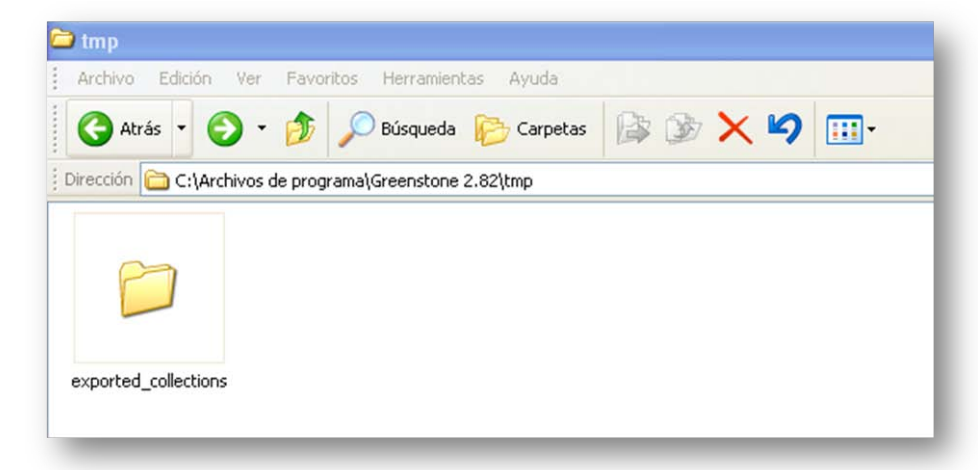


Figura 17.6. Colección exportada en carpeta tmp.

Para completar el proceso se utiliza un quemador de CD/DVD para copiar sus contenidos en el CD en blanco, de esta forma usted puede consultarla en cualquier computadora u ordenador sin que para ello esté instalado el software Greenstone.

El resultado será un CD-ROM o un DVD Greenstone autoinstalable, el cual inicia el proceso de instalación tan pronto como se haya colocado el disco. Si desea explorar la colección exportada puede ver su estructura interna desde el disco.

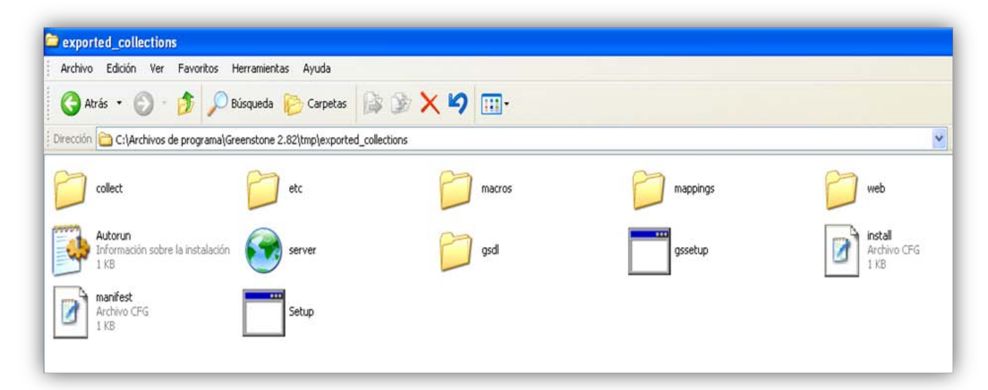


Figura 17.7 Estructura interna de la carpeta exportada en el CD-ROM.

La figura anterior presenta la colección exportada con su estructura interna. Es con el icono de server (en verde y azul) que al darle un clic se visualiza inmediatamente la colección tal y como se muestra en la imagen siguiente

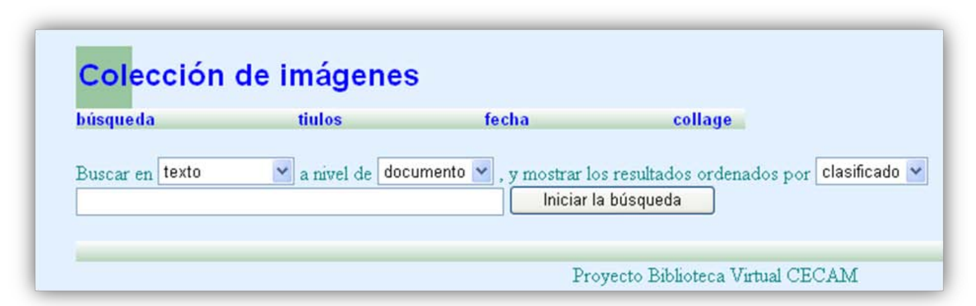


Figura 17.8 Interfaz de la colección exportada.

Esta pantalla visualiza la interfaz de la colección de imágenes con los clasificadores escogidos durante el diseño y configuración de la colección, en este caso títulos, descripción, collage (Ver Figura 17.9)



Figura 17.9 Interfaz de la colección de imágenes con clasificadores seleccionados.

En la figura está activado el clasificador Títulos por el cual el usuario va a consultar la información, se aprecian además los clasificadores Descripción y Collage.

Una vez elegida la búsqueda por los clasificadores títulos, muestra además la información y descripción, resultado de la forma en que se concibieron las cadenas de formato durante la creación y configuración de la colección.

17.4- Consideraciones Finales.

La opción exportar a CD es una de las bondades mas importantes que exhibe el programa Greenstone que permite visualizar las colecciones creadas aun en computadoras en las que no este instalado el programa, lo que representa una gran ventaja para la difusión y socialización del conocimiento.

17.5- Referencias Bibliográficas

(1) Tramullas J. Biblioteca Digital Greenstone. Capitulo 10. En.Software Libre para Servicios de Información [Libro en Internet] Digital. Madrid: Pearson Prentiss Hall, 2006. Disponible en:<http://tramullas.com>

(2) Miranda C. Sistemas de Gestión de Documentos bajo Licencia de Código Libre [Sitio Web en Internet] 08 Junio 2009. Disponible en:<http://opengraphprotocol.org/schema//xmlns:fb>

(3) Nogales J.T. MARC en XML [Sitio Web en Internet] 2004-04-28. Disponible en:<http://www.bib.uc3m.es/~nogales/cursos/marcxml.html=intro>

(4) Library of Congress. Metadata Encoding and Transmission Standard (METS) Official Web Site [Website in Internet], November 25, 2009. Available from: <http://www.loc.gov/standards/mets/mets-home.html>

(5) Grupo de usuarios Greenstone de Latinoamérica .Greenstone v2.82 [Sitio Web en Internet]

2008.Disponible en < http://www.greenstone-la.org/descarga-de-software/57-gs-282>

(6) Bainbridge D, McKay D, Witten H. I. Guía del programador. Biblioteca digital Greenstone. [Sitio Web en Internet].Departamento de Informática. Universidad de Waikato, Nueva Zelanda. Marzo de 2004.Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Develop_es_all.html>

CAPÍTULO 18

18- Términos Conceptuales y Referenciales

Dra. C. Caridad Fresno Chávez (cfresno@cecam.sld.cu). Prof. Titular, MSc Gestión de Información. Centro de Cibernética Aplicada a la Medicina (CECAM) Dirección: Calle 146 # 2511 Esq.31 Playa, Ciudad Habana, Cuba, CP: 11600, Telf.: (537) 271-1354. Universidad de Ciencias Médicas de La Habana (UCMH)

Tec. **Dailyn Rodríguez Martínez**. Técnico Medio en Informática. Estudiante de 5to año de la carrera de Bibliotecología y Ciencia de la Información.

Alfabetización Digital:

Basada en las tecnologías de redes se concibe como un subconjunto de la alfabetización general, ya que incluye los aspectos relacionados con el conocimiento y uso de los servicios de información electrónica, así como la comprensión de los sistemas de información a través de los que se accede a los contenidos.

Bawden D. Revisión de los conceptos de alfabetización informacional y alfabetización digital. Ann Document. 2002;5:361-408.

Alfabetización Informacional:

1-Significa incorporar capacidades para el dominio de las nuevas tecnologías de información y comunicaciones que contribuyan al desarrollo de los recursos humanos o capital humano como única vía o alternativa economica para alcanzar un verdadero desarrollo sostenible, en las economías de países desarrollados o no.

UNESCO Communication, Information, Informatics Sector (CII). Creating a new UNESCO programme for a just and free information society [Website in Internet] Paris. UNESCO, 2000. Disponible en: <http://www.unesco.org/webworld/future/introduction.shtml>.

2-se entiende la alfabetización en información como un conjunto de aptitudes para localizar, manejar y utilizar la información de forma eficaz para una gran variedad de finalidades, basada en; las tecnologías de la información las fuentes de información la información como proceso elcontrol de la información la construcción de conocimiento la extensión del conocimiento.

Bruce S. Las siete caras de la alfabetización en información en la Enseñanza Superior. Anales de Documentación [Seriada en Internet]. 2003, No.6, Disponible en: <http://www.um.es/fccd/anales/ad06/ad0600.html> Traducción del trabajo original: Bruce S. Seven faces of information literacy in Higher Education [Website in Internet].1997. Available from: <http://sky.fit.qut.edu.au/~bruce/inflit/faces/faces1.htm>

Archivo *main.cfg:*

Contiene información común a la interfaz de todas las colecciones presentadas en el sitio Greenstone. Incluye la dirección de correo electrónico del encargado del mantenimiento del

sistema, indica si las páginas de estado y del Colector están activadas o no, si se llevan o no los registros de las actividades de los usuarios y si se utilizan o no *"cookies"* para identificar a los usuarios.

Greenstone .org. Biblioteca Digital Greenstone. Guía del Programador .Configuración del sitio Greenstone. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm>

Archivo de configuración principal main.cfg:

Sirve para configurar el recepcionista, esto es, la parte de Greenstone que procesa las consultas y muestra las páginas. Se puede controlar todo, desde las lenguas que la interfaz puede utilizar hasta los registros que se llevarán. El archivo de configuración *main.cfg* contiene dos tipos de entradas que inciden en la gestión de las diferentes lenguas. Éstas determinan qué idiomas y qué códigos estarán disponibles en la página de preferencias.

Greenstone.org. Biblioteca Digital Greenstone. Guía del Programador .Configuración del sitio Greenstone. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm>

Arquitectura de la Información:

1. El término "Arquitectura de la Información" (AI) fue utilizado por primera vez por Richard Saul Wurman en 1975, quién la define como: El estudio de la organización de la información con el objetivo de permitir al usuario encontrar su vía de navegación hacia el conocimiento y la comprensión de la información.

   Wurman R 1975. Citado por Martín F.J. Hassan Y. Qué es la Arquitectura de la Información. No Solo Usabilidad [Seriada en Internet] no. 2, 2003/02/16. Disponible en: <http://www.nosolousabilidad.com/hassan>

2. El arte y la ciencia de estructurar y clasificar sitios Web e intranets con el fin de ayudar a los usuarios a encontrar y manejar la información.

   Rosenfeld l. Morvillr P.Information Architecture for the World Wide Web: Designing Large-Scale Web Sites,[Wesite in Internet] 2nd. August 15, 2002. Disponible en: <Louis Rosenfeld (Author) <http://www.amazon.com/Louis-Rosenfeld/e/B000APV7C0/ref=ntt_athr_dp_pel_pop_1> Citado por: Martín F.J. Hassan Y. Qué es la Arquitectura de la Información. No Solo Usabilidad [Seriada en Internet] no. 2, 2003/02/16. Disponible en: <http://www.nosolousabilidad.com/hassan>

Base de datos MySQL:

1. MySQL es un sistema de gestión de bases de datos relacional, licenciado bajo la GPL de la GNU. Su diseño multihilo le permite soportar una gran carga de forma muy eficiente. Las principales características de este gestor de bases de datos son las siguientes:
   1. Aprovecha la potencia de sistemas multiprocesador, gracias a su implementación multihilo.
   2. Soporta gran cantidad de tipos de datos para las columnas.
   3. Dispone de API's en gran cantidad de lenguajes (C, C++, Java, PHP, etc).
   4. Gran portabilidad entre sistemas.
   5. Soporta hasta 32 índices por tabla.
   6. Gestión de usuarios y passwords, manteniendo un muy buen nivel de seguridad en los datos.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

Pecos D. 3-MySQL.PostGreSQL vs. MySQL. Manual.[Sitio Web en Internet] Disponible en:<http://danielpecos.com/docs/mysql_postgres/b164.html#MYSQL>

2. MySQL o My Structured Query Language o Lenguaje de Consulta Estructurado, se remite a principios de la década de 1980. Programadores de IBM lo desarrollaron para contar con un código de programación que permitiera generar múltiples y extendidas bases de datos para empresas y organizaciones de diferente tipo. Una de las características más interesantes de MySQL es que permite recurrir a bases de datos multiusuario a través de la Web y en diferentes lenguajes de programación que se adaptan a diferentes necesidades y requerimientos. Por otro lado, MySQL es conocida por desarrollar alta velocidad en la búsqueda de datos e información, a diferencia de sistemas anteriores. Las plataformas que utiliza son de variado tipo y entre ellas podemos mencionar LAMP, MAMP, SAMP, BAMP y WAMP (aplicables a Mac, Windows, Linux, BSD, Open Solaris, Perl y Phyton entre otras).

   Definición ABC Definición de MySQL [Sitio Web en Internet] 2011-09-12 .Disponible en:<http://www.definicionabc.com/wp-content/themes/definicion/style.css> y <http://www.mysql.com/>

Biblioteca Digital:

1. Christine Borgman menciona en 1999 dos enfoques distintos que son utilizados entre los profesionales de las bibliotecas. Ella describe "un conjunto de recursos y de capacidades técnicas para crear, buscar y utilizar la información", pero subraya también que "las bibliotecas digitales están concebidas y organizadas para (y por) una comunidad de usuarios, y sus medios son utilizados para responder a las necesidades de información y a los usos de esta comunidad".

   Le Crosnier H. Bibliotecas Digitales, Citado en: Palabras en Juego: Enfoques Multiculturales sobre las Sociedades de la Información.[Libro en Internet] Editores: Alain Ambrosi, Valérie Peugeot y Daniel Pimienta. 5 de noviembre de 2005. Editorial por C&F Éditions. . Disponible en:< http://cfeditions.com/> Disponible en: <http://vecam.org/skel/css/edm/style_print.css" media=print>

2. Las ventajas de las bibliotecas digitales pasan por permitir un acceso universal a información estructurada y ordenada, pero para ello es necesario la creación y seguimiento de estándares, el diseño de una buena estructura capaz de soportar los servicios, así como establecer una política de acceso, los derechos de autor y la legislación sobre propiedad intelectual. Todos son aspectos de vital importancia para la creación, como para la protección de estas bibliotecas.

   Red de Investigación en Bibliotecas Digitales. Catálogos de Bibliotecas Virtuales[Libro en Internet] 2001.Disponible en: <http://www.jbidi.org>

3. Las bibliotecas pueden clasificarse en tres tipos:

   - 1) Analógicas/Bibliotecas de Papel (BP) - la biblioteca de papel clásica con su fichero manual.
   - 2) Automatizada/Biblioteca Híbrida (BH) - una biblioteca analógica con el catálogo automatizado.
   - 3) Biblioteca Digital (BD) - una biblioteca automatizada en la que la mayor parte de la información es digital.

Sharon T, Frank A J Bibliotecas digitales en Internet - 66th IFLA Council and General

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

Conference - Conference Programme and Proceedings. [Website in Internet] 2000-08-22 Available from: <http://www.ifla.org/>new website>

Brecha Digital:

Es el producto de las brechas sociales y económicas existentes entre los continentes, países estados, municipios y sociedades del mundo.

Gaynor M. Brecha digital- ¿un problema? La Brecha Digital. [Sitio Web en Internet] Artículos 2002. Disponible en: <http://www.labrechadigital.org/articulos.html>

C++:

1. El C++ es un lenguaje de programación, diseñado a mediados de los años 1980, por Bjarne Stroustrup, como extensión del lenguaje de programación C. Se puede decir que C++ es un lenguaje que abarca tres paradigmas de la programación: la programación estructurada, la programación genérica y la programación orientada a objetos.

   PergaminoVirtual.Glosario [Sitio Web en Internet] 2007-06-12 , Disponible en: <http://www.pergaminovirtual.com.ar/definicion/default.css>

2. Lenguaje de programación en que está escrita la mayor parte del programa Greenstone.

   Greenstone Wiki. Biblioteca Digital Greenstone. Guía de Usuario. Apéndice B: Glosario de Términos [Sitio Web en Internet], Marzo 2006. Disponible en: <http://greenstonewiki.cs.waikato.ac.nz/wiki/gsdoc/GNUFDL.html>

3. C++ se ha popularizado porque combina la programación tradicional en C con programación orientada a objetos. Smalltalk y otros lenguajes originales de programación orientada a objetos no suministraban las estructuras familiares de lenguajes convencionales como C y Pascal.

   Mastermagazine. Definición de C++.[Sitio Web en Internet] 10 de febrero 2005. Disponible en: <http://www.mastermagazine.info>

Cadenas de Formato:

Las páginas Web que usted ve cuando utiliza Greenstone no han sido almacenadas previamente sino que son generadas “al instante” a medida que se las necesita. La apariencia de numerosos aspectos de las páginas se controla a través de las “cadenas de formato”. Las cadenas de formato se encuentran en el archivo de configuración de la colección y se introducen con la palabra clave format seguida del nombre del elemento al que se aplica el formato. Las cadenas de formato controlan dos tipos diferentes de elementos de página. El primero comprende los elementos de la página que muestran documentos o partes de documentos. El segundo incluye las listas generadas por los clasificadores y las búsquedas. Todas las cadenas de formato se interpretan en el momento en que se visualizan las páginas. Es rápido y sencillo experimentar con las cadenas de formato ya que surten efecto en cuanto se guarda cualquier modificación en el archivo collect.cfg. Ejemplo de cadena de formato: “*Expand Text|Expand contents|Detach|Highlight*”. Los archivos de macros tienen la extensión .dm. Cada archivo define uno o más paquetes (packages), cada uno de los cuales contiene una serie de macros utilizadas con un fin preciso y único. Al igual que los clasificadores y los conectores, hay una base a partir de la cual se elaboran las macros, que se llama base.dm; este archivo define el contenido básico de una página. Los nombres de las macros empiezan y terminan con una rayita de subrayado y su contenido se define mediante llaves. El contenido puede ser texto llano, HTML (incluidos los enlaces con pequeños

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

programas (applets) de Java y JavaScript), nombres de macros o cualquier combinación de estos elementos.

Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 2. Explotación Óptima de sus Documentos. [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

Calidad de la Información:

Se trata de la calidad de las comunicaciones de negocios en todas las formas –hablada, escrita, correo electrónico, reportes, Internet, Intranet, manuales de política, catálogos, newsletters, así como las bases de datos y dataware house. Pero más que eso, se trata de servir al cliente, proveyéndolo de productos de información y servicios que "consistentemente cumplan las expectativas del cliente y de los trabajadores del conocimiento".

English L. Information Quality in the Knowledge Age. [Serials in Internet] Resource Portal: Portal Related Articles. DM Review, October 1999 : Available from: <http://www.dmreview.com>

CDS/ISIS:

CDS/ISIS permite construir y manejar bases de datos estructuradas no numéricas, es decir bases de datos constituidas en su mayor parte por texto. Si bien CDS/ISIS trabaja con textos y palabras, y ofrece por tanto muchas de las herramientas normalmente encontradas en los paquetes de procesadores de textos, no consiste exactamente en un procesador de textos. Esto es así porque los textos que CDS/ISIS procesa están estructurados en elementos de datos definidos por el usuario. CDS/ISIS es un sistema generalizado de almacenamiento y recuperación de información manejado por un conjunto de menús diseñado específicamente para el manejo por ordenador de bases de datos estructuradas. Una de las mayores ventajas ofrecidas por el diseño del sistema consiste en que CDS/ISIS es capaz de manipular un número ilimitado de bases de datos cada una de las cuales puede consistir en un conjunto diferente de datos elementales.

Del Bigio G. CDS/ISIS for Windows. Manual de Referencia (Versión1.31). UNESCO, Noviembre 1998. Traducido por Adolfo Hernández, CINDOC-CSIC. Distribuidor Nacional para España

CGI (Common Gateway Interface- (interfaz común de Pasarela):

1. Sistema que permite que los usuarios puedan activar programas en la computadora central (o servidor), haciendo clic en páginas Web.

   Greenstone.org. Guía de Instalación Greenstone. [Sitio Web en Internet]. 2006. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone>

2. Interface de Acceso Común. Estos son programas usados para hacer llamadas a rutinas o bien para controlar otros programas como también bases de datos desde una página Web. También tienen la habilidad de generar directamente HTML, o sea crear una página Web con información. Tiene también la habilidad (Interfaz) de intercambiar datos estándar en WWW organizando el envío y recepción de datos entre el navegador y programas que residen en el servidor a través del cual se organiza el envío y recepción de datos entre navegador y programas residentes en servidores WWW. Los CGIs pueden programarse en Perl y C.

Plantillas Web. Definición de cgi - Common Gateway Interface! v1.1. [Sitio Web en Internet]. Nov 28, 2005. Disponible en: <http://www.descargas.walkever.com>

3. Los lenguajes más comunes para escribir estos programas son el lenguaje perl, C ó bien un script escrito en cualquier shell (csh, sh, ksh ó zsh). Algo importante es que los cgi-bin deben tener una extensión especial .cgi, o bien, encontarse en un directorio especial llamado cgi-bin. Cuando se referencie en una pagina html un binario, cualquier fichero con la extensión cgi, en el servidor se ejecutará este programa y la salida se enviará al cliente de la WWW. Para que sea posible ejecutar el binario (cgi-bin), será necesario que el fichero tenga los permisos de forma correcta, para ello tendrá que ejecutar en Unix el siguiente comando:

   CGI (Common Gateway Interface), ejecución de binarios. [Sitio Web en Internet]. 2002. Disponible en: <http:www2.uca.es/manual-html/cgi.bin> 14-cgi-bin.

4. Es un directorio desde el cual el servidor Web trata documentos como programas ejecutables. La directiva ScriptAlias de Apache se utiliza para crear un directorio cgi-bin. Cabe señalar que esta directiva puede hacer de cualquier directorio un directorio CGI ejecutable, sin que se llame forzosamente "cgi-bin" A la inversa, un directorio llamado "cgi-bin" no tiene nada especial si no se le ha aplicado la directiva ScriptAlias.

   Greenstone.org. Guía de Instalación Greenstone. [Sitio Web en Internet]. 2006. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone>

cgi bin:

Es un directorio desde el cual el servidor Web trata documentos como programas ejecutables. La directiva ScriptAlias de Apache se utiliza para crear un directorio cgi-bin. Cabe señalar que esta directiva puede hacer de cualquier directorio un directorio CGI ejecutable, sin que se llame forzosamente "cgi-bin" A la inversa, un directorio llamado "cgi-bin" no tiene nada especial si no se le ha aplicado la directiva ScriptAlias.

Greenstone.org. Guía de Instalación Greenstone. [Sitio Web en Internet]. 2006. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone>it

Clasificadores en Greenstone:

Los clasificadores sirven para crear los índices de consulta de la colección Los clasificadores, al igual que los conectores, se especifican en el archivo de configuración de la colección. Todos tienen una línea que empieza por la palabra clave classify seguida del nombre del clasificador y las opciones que admite. El clasificador general es DateList, por ejemplo genera una lista de selección de intervalos de fechas. Otros clasificadores generan estructuras de consulta que son explícitamente jerárquicas. Las clasificaciones jerárquicas son útiles para las clasificaciones y subclasificaciones temáticas, así como para las jerarquías relativas a organizaciones. Todos los clasificadores generan una estructura jerárquica que se utiliza para mostrar un índice de consulta. Los clasificadores de Greenstone son: Hierarchy, List, SectionList , AZList , AZSectionList , DateList

Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 2. Explotación Óptima de sus Documentos. [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

db2txt/:

Convierte el formato de base de datos de GNU en formato de texto ASCII de tipo XML.

Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 2. Explotación Óptima de sus Documentos. [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

Derecho de Autor:

1. Queda definido como el conjunto de normas y principios que regulan los llamados derechos patrimoniales que la ley concede a los autores; por el solo hecho de la creación de una obra ya sea literaria, artística, o científica, tanto publicada o que todavía no se haya publicado.

   Boutang Y. Nuevas tecnologías de la información y la comunicación. Enfoque económico. Boletín del Instituto de Investigaciones Económicas. [Seriada en Internet]. 11 de febrero de 2004. Disponible en: <http://www.iiec.unam.mx/notiiiec/extras/conferencia_Yann_Moulier_Boutang.pdf>

2. Se entiende la protección jurídica que se otorga al titular del derecho de una obra original del que es autor. El derecho de autor comprende las categorías principales de derechos: los derechos patrimoniales y los derechos morales .

   OMPI Gestión Colectiva del Derecho de Autor y los Derechos Conexos. [Sitio Web en Internet] Documento de Posición. 2000. Disponible en <http://www.wipo.org/about-ip/es/collective-mngt.html>

Desarrollo Sostenible:

La actual «crisis de mutación» del capitalismo impone el paso hacia un concepto de desarrollo sostenible concebido como la constitución de una sociedad de la «democracia y de la cooperación de los saberes», en la que, para decirlo en las palabras de K. Marx en los Grundrisse, el «principal capital fijo pasa a ser el hombre mismo». El Desarrollo Sostenible responde a:

1. Integración de la Conservación y el Desarrollo.
2. Satisfacción de las necesidades Básicas Humanas.
3. Provisión para la autodeterminación social y la diversidad cultural, humana y geográfica.
4. Mantenimiento de la integridad ecológica.

UN. Naciones Unidas División de Desarrollo Sostenible - Programa 21 - Capitulo 4 [Sitio Web en Internet]24 March 2003. Disponible en:<http://www.un.org/esa/sustdev/documents/WSSD_POI_PD/Spanish/WSSDsp_PD.htm >

DSpace:

1. DSpace es un software de código abierto diseñado por el Massachusetts Institute of Technology (MIT) y los laboratorios de HP para gestionar repositorios de ficheros (textuales, audio, vídeo, etc.), facilitando su depósito, organizándolos en comunidades, asignándoles metadatos y permitiendo su difusión a recolectores o agregadores. DSpace se una aplicación cliente/servidor que se gestiona vía Web, es decir, que la mayor parte de procesos pueden llevarse a cabo con un navegador estándar como Internet Explorer, Firefox u Opera. DSpace estructura los repositorios en comunidades y colecciones dónde las comunidades contienen sub-comunidades y/o colecciones y las colecciones contienen

documentos.

Rodríguez-Gairín J.M, Sulé A. DSpace: un manual específico para gestores de la información y la documentación.[Libro en Internet].Junio de 2008. Disponible en: <http://www.ub.edu/bid/20rodri2.htm>

2. La combinación de requerimientos de dspace es un poco inusual y antes de poder instalarlo y habilitarlo es necesario instalar y configurar Sun Java, Apache Tomcat, Apache Ant, Apache Maven, PostgreSQL y Perl. Lamentablemente, Dspace no proporciona un paquete integrado o proceso de instalación automatizado que simplifique el proceso y permita instalar la aplicación y los software requeridos cómodamente.

   Miranda C. Sistemas de Gestión de Documentos bajo Licencia de Código Libre [Sitio Web en Internet] 08 Junio 2009. Disponible en:<http://opengraphprotocol.org/schema//xmlns:fb>

Documento:

Un documento es la prueba o testimonio material de un hecho o acto que una persona física o jurídica, una institución, asociación, etc., que puede ser de carácter público o privado, realizan como consecuencia del ejercicio de sus actividades y funciones y que podrá ser plasmado en una unidad de información que observe cualquier soporte, papel, cinta, disco magnético, película y fotografía, con el objeto de preservarlo en el tiempo en caso de necesitarlo para presentarlo como prueba, recuerdo o legado a alguien. Un documento está compuesto preeminentemente por información escrita que puede haber sido volcada de puño y letra, o sea, manuscritamente, o bien a través de un proceso mecánico como puede ser el de la máquina de escribir que se usaba masivamente antaño o por la computadora, que es aquel medio mecánico que predomina hoy en día.

DefinicionABC. Definición de Documento,[Sitio Web en Internet]. 2006. Disponible en: <http://www.definicionabc.com>

Economía del Conocimiento:

Término que describe las bases para la creación del valor en la actual economía global. El foco central de esta economía, es el conocimiento a través de la innovación, más que a través de la producción y distribución masiva. La explotación inteligente de la información, su conversión en conocimiento, la generación de mecanismos de aplicación de ese conocimiento en el desarrollo de nuevos productos y procesos, incluso de nuevas formas de organización social, es, pues, el recurso fundamental, en la economía del conocimiento

Stewart T. Intellectual Capital: The New Wealth of Organization. Doubleday/Currency. N.Y. Tom Stewart on Intellectual Capital. Interview with Lewis Perelman [Website in Iinternet].KI Executive Editor, 1999: Available from: <http://members.aol.com/thosstew/> y <http://www.knowledgeinc.com>

Educación a distancia:

1. Puede definirse como la acción o proceso de educar o ser educado, cuando este proceso se realiza a distancia. Sangra Albert. Educación a distancia, educación presencial y usos de la tecnología: una triada para el progreso educativo.

   Edutec. Revista Electrónica de Tecnología Educativa.[Seriada en Internet] (15), 2002. Disponible en:<http://www.unib.es/depart/gte/edutec_e/recelec15/albert_sangra.htm>

2. La Educación a distancia (EAD) se define de modo mas acucioso como un sistema tecnológico de comunicación masiva y bidireccional, que sustituye la interacción personal

profesor-estudiante en el aula como medio preferente de la enseñanza por la acción sistemática y conjunta de diversos recursos didácticos y el apoyo de una organización tutorial. En esta definición se pueden precisar los siguientes rasgos:

- La separación del profesor y el alumno durante la mayor parte del proceso de instrucción.
- La influencia de un organismo o institución educacional a través de la tutoría y la evaluación.
- El uso de tecnologías para la transportación de los contenidos del curso.
- Existencia de comunicación bidireccional entre el profesor, el tutor o la institución educativa y el estudiante.

  Simonetti, G., Drogo G. El desafío de formación de recursos humanos en enfermería con modalidad de educación a distancia. Rev. de Enfermería del Hospital Italiano. [Seriada en Internet]. 2 (6). Disponible en: <http://www.hitalba.edu.ar/docencia/rehi/index.php?contenido=contenido_articulo.php&id_articulo=544 >

Estándares:

1. Común y repetido uso de reglas, condiciones, lineamientos o características de los productos o procesos relacionados y producción de métodos, y manejos relacionados en la práctica sistemática.

   Library of Congress. Technical Reports and Standards [Website in Internet]. 2005 Available from: <http://www.loc.gov/global/stdLauncher.js >

2. Existen varios estándares y lineamientos que son importantes para así poder garantizar un valor presente y futuro de los documentos gestionados y su información, su expansión, integración e interconexión con o migración a otros sistemas. Estos estándares incluyen el Protocolo para Cosecha de Metadata (Protocol for Metadata Harvesting: OAI-PMH) de la Open Archives Initiative, así como OAIS, el modelo de referencia ISO para Open Archival Information System, definido por recomendación del Comité Consultivo para Sistemas de Datos Espaciales (CCSDS) y el Dublin Core para definición y estructura de metadatos, entre otros.

   Miranda C. Sistemas de Gestión de Documentos bajo Licencia de Código Libre [Sitio Web en Internet] 08 Junio 2009. Disponible en:<http://opengraphprotocol.org/schema/xmlns:fb>

Estándar Dspace:

Por defecto, los formularios que se utilizan para describir los documentos son iguales para todas las colecciones y están basados en el esquema Dublin Core. Sin embargo, editando el fichero "input-forms.xml" se pueden crear formularios específicos para colecciones, así como asignar un nuevo esquema de metadatos para la descripción del recurso. También desde este fichero se puede especificar en cada campo si es o no obligatorio, si es o no repetible y asignarle un lenguaje controlado para la introducción de los valores.

Miranda C. Sistemas de Gestión de Documentos bajo Licencia de Código Libre [Sitio Web en Internet] 08 Junio 2009. Disponible en:<http://opengraphprotocol.org/schema/xmlns:fb >

Estándar MARCXML:

Iniciativas de la Network Development and MARC Standards Office (Oficina de Desarrollo de Redes y Normas MARC) de la Biblioteca del Congreso de Estados Unidos para la conversión de registros de catalogación MARC a marcado SGML (Standard Generalized Markup Language) o XML (Extensible Markup Language) y viceversa sin pérdida de datos.

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

Estas iniciativas han originado estructuras de datos en forma de DTD (Document Type Definition, Definición de Tipo de Documento) o Esquemas XML, herramientas de conversión, manuales y hojas de estilo. [V. MARC en XML, en http://www.loc.gov/marc/marcxmlspa.html]

- MARC SGML (1995),
- MARC XML DTD (2001).
- MARCXML Schema (2002)

  Las DTD son de gran tamaño, ya que cada elemento de datos MARC (campo o subcampo) es un elemento XML distinto. Para simplificar la estructura y mejorar la integración con aplicaciones XML orientadas a datos se ha desarrollado un Esquema XML (v. 1.0 2002-05-21, v. 1.1 2003-08-04). [V. MARCXML : MARC 21 XML Schema http://www.loc.gov/standards/marcxml].

Nogales J.T. MARC en XML [Sitio Web en Internet]2004-04-28. Disponible en: en:<http://www.bib.uc3m.es/~nogales/cursos/marcxml.html=intro>

Estrategia de Búsqueda:

1. Para conseguir llevar a cabo una buena tarea de búsqueda son necesarios tres ingredientes fundamentales: diseñar un perfil de búsqueda adecuado; es decir hay que tener la habilidad de emparejar con exactitud los términos que se utilizan en el perfil de búsqueda con los que se encuentran en los documentos cuya información se quiere encontrar; cononcer el tamaño y contenido de la base de datos que se elige, y conocer la posibilidades que ofrecen las diferentes herramientas de búsqueda para recuperar los contenidos de su base de datos.

   Universitas Miguel Hernández. Buscando información en Internet. Estrategia de búsqueda en Internet. [Sitio Web en Internet]10 de julio de 2003. Disponible en: <http://www.dsp.umh.es/docent/ienol/estrategia.htm>

2. Una buena estrategia de búsqueda de información seria la siguiente: identifique los conceptos clave y el área a la que pertenece el objeto de su búsqueda. La clase de fuente donde puede encontrarse, por ejemplo: Web, Usenet, Gopher u otros ... Primero seleccionar un buscador temático para ver la información genérica. Después use un buscador automático para obtener información más específica. Terminar con buscadores especializados.

   Vicent A. Figueras F.. La manera más eficiente de buscar información en Internet. 2002

Gestión del Conocimiento:

1. Involucra la identificación y análisis del conocimiento disponible y los procesos relacionados así como el control en las acciones a desarrollar para cumplir los objetivos Organizacionales.

   Macintosh A. Knowledge Management. AIAI Artificial Intelligence Applications Institute. University of Edinburgh, [Website in Internet] January 1999. Available from: <http://aiai.ed.ac.uk/967alm/kamlnks.html>

2. Se considera el arte de la creación de valor a partir de activos intangibles.

   Sveiby K Intellectual Capital and Knowledge Management. , [Website in Internet] April 1998. Available from: http://www.sveiby.com.au/KnowledgeManagement.html

Ghostscript:

1.- Ghostscript (escrito por Peter Deutsch, fundador de Aladdin Enterprises), es el programa intérprete por excelencia de documentos en formato PS (y también PDF). Lenguaje Postscript que fue desarrollado por la empresa Adobe, como formato de representación gráfica. La idea consiste en que una imágen, compuesta por figuras geométricas y por texto en tipos de letras y tamaños arbitrario puede ser representada de forma descriptiva. La formalización de esta descripción forma un lenguaje de programación.

Indaganda.com - Definición de Ghostscript y Información sobre Ghostscripthttp [Sitio Web en Intenet] 2001. Disponible en: <http://www.indaganda.com/definir/Ghostscript#>

2.-Actualmente casi todas las impresoras permiten imprimir documentos PS (postscript) pero algunas no disponen de esa opción, para ello se emplea este software permitiendo extraer texto de documentos con formato "ps, eps, pdf", el modo de impresión, conversión de un formato a otro, etc. Para obtener mas información de estos programas se recomienda visitar la siguiente página Web: http://www.cs.wisc.edu/~ghost

UNED - Centro de Servicios Informáticos . Ghostscript & Gswiew[Sitio Web en Internet]2009. Disponible en:<http://www.uned.es/csi/estilo2.css" type=text/css rel=stylesheet>

GLI (Greenstone Library Interfaz):

1.- La interfaz del bibliotecario de Greenstone es una aplicación java, de entorno grafico, que se incorpora a partir de la versión 2.41. Desarrolla con detalle las opciones de configuración y facilita el trabajo de configuración de la colección. Se ejecuta de modo local, aunque en las últimas versiones de Greenstone se ha desarrollado un applet de java que realiza las mismas funciones a través del servidor Web lo que permite el trabajo remoto.

Tramullas J. Biblioteca Digital Greenstone. Capitulo 10. En.Software Libre para Servicios de Información[Libro en Internet] Digital. Madrid: Pearson Prentiss Hall,2008. Disponible en:<http://tramullas.com>

2.- La interfaz grafica de Greenstone es una herramienta de asistencia a los que elaboran las colecciones con este software. Se instala automáticamente con todas las versiones de Greenstone , Se coloca en el subdirectorio gli del nivel superior del Directorio Greenstone.

Greenstone.org. Guía de Instalación Greenstone. [Sitio Web en Internet]. 2002. Disponible en:
<http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone>

Globalización:

El proceso de globalización se puede definir como el ensanchamiento geográfico de los nexos internos de productos, mercados, empresas y factores de producción con un componente mayor del cual fueron derivados, generados o disponibles en otros países y regiones OECD, 1992. Representa un proceso de interconexión financiera, económica, política y cultural que se relaciona con la progresiva expansión internacional de las actividades empresariales gestadas a partir de la II Guerra Mundial que cobra fuerza desde 1970 con la crisis económica que promueve la incorporación de la Nuevas Tecnologías Comunicacionales

Campo C. La globalización como proceso social ambiguo. [Sitio Web en Internet].Tips. Sistema de Promoción de Información Tecnológica y Comercial. 2006. Disponible en:

<http://www.redtips.org/tips/forum/Sinformacion/0006.htm>

GSDLHOME:

Los elementos definidos en el directorio GSDLHOME/lib son objetos Greenstone de nivel inferior, creados por encima de STL, y que se utilizan en todo el código fuente. Así por ejemplo, el código fuente del sistema de ejecución se encuentra en GSDLHOME/src. Ocupa dos subdirectorios, recpt para el código del recepcionista y colservr para el del servidor de colección. Otro directorio, GSDLHOME/lib, incluye objetos de nivel inferior que el recepcionista y el servidor de colección utilizan. Los archivos de encabezados situados en el directorio GSDLHOME/lib comprenden una mezcla de funciones y objetos que contribuyen útilmente al sistema de ejecución de Greenstone. Cuando lo importante es la eficiencia, las funciones y las funciones miembro se especifican inline. En la mayoría de los casos, los detalles de implementación se incluyen en el archivo .cpp correspondiente al archivo de encabezados.

Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3. El Sistema de Ejecución de Greenstone [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

Greenstone:

Es un completo sistema diseñado para construir y organizar colecciones de recursos digitales, creado por el Proyecto d e Biblioteca Digital de Nueva Zelanda con sede en la Universidad de Waikato , es desarrollado y distribuido por la UNESCO con la colaboración de la ONG "Información para el Desarrollo Humano" con sede en Bélgica. Geenstone es un software open source (o software de código abierto)que se distribuye de acuerdo a los términos de la Licencia Publica General (GNU) http://www.gnu.org/gnu/gnu-history.es.html Existen sendas versiones de Greenstone disponibles para Windows y Unix, como binarios y en forma de código fuente.

Greenstone .org. Biblioteca Digital Greenstone. Guía del Programador .Configuración del sitio Greenstone. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm >

Guión CGI:

1.- El CGI por sus siglas en inglés „Common Gateway Interface" es de las primeras formas de programación Web dinámica. En sí, es un método para la transmisión de información hacia un compilador instalado en el servidor. Su función principal es la de añadir una mayor interacción a los documentos Web que por medio del HTML se presentan de forma estática.

El CGI es utilizado comúnmente para contadores, bases de datos, motores de búsqueda, formularios, generadores de email automático, comercio electrónico, rotadores y mapas de imágenes, juegos en línea y otros. Esta tecnología tiene la ventaja de correr en el servidor cuando el usuario lo solicita por lo que es dependiente del servidor y no de la computadora del usuario.

Van Der C. ¿Qué es el CGI? [Sitio Web en Internet] 23 mar 2001. Disponible en: <http:www.maestrosdelweb.com >

2.- Código asociado a un botón, un menú, o un enlace en una página Web, que especifica lo que ha de hacer la computadora central cuando se hace clic en ese objeto.

Greenstone.org. Guía de Instalación Greenstone. [Sitio Web en Internet]. Disponible en:

<http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone>

Hash:

Hash se refiere a una función o método para generar claves o llaves que representen de manera univoca a unn documento, registro, archivo, resumir o identificar un dato a través de la probabilidad, utilizando una función hash o algoritmohash

Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3. El Sistema de Ejecución de Greenstone [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

hashfile:

Programa autónomo incluido en Greestone que calcula un identificador de documento único, basándose en el contenido del archivo. Cada parte del documento tiene asociado un hash que es un algoritmo único para cada documento o parte del documento. El sistema asigna un sistema alfanumérico único que nunca se podrá utilizar aunque se elimine de forma lógica.

Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3. El Sistema de Ejecución de Greenstone [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

Hojas de Estilo en Cascada (Cascading Styles Sheets-CSS)

1.-CSS son las siglas de Cascading Style Sheets - Hojas de Estilo en Cascada - que es un lenguaje que describe la presentación de los documentos estructurados en hojas de estilo para diferentes métodos de interpretación, es decir, describe como se va a mostrar un documento en pantalla, por impresora, por voz (cuando la información es pronunciada a través de un dispositivo de lectura) o en dispositivos táctiles basados en Braille. CSS es una especificación desarrollada por el W3C (World Wide Web Consortium) para permitir la separación de los contenidos de los documentos escritos en HTML, XML, XHTML, SVG, o XUL de la presentacin del documento con las hojas de estilo, incluyendo elementos tales como los colores, fondos, márgenes, bordes, tipos de letra..., modificando as la apariencia de una página Web de una forma más sencilla, permitiendo a los desarrolladores controlar el estilo y formato de sus documentos.

W3C.Hojas de estilo o cascading style sheets [Sitio Web en Internet] 2005. Disponible en: <www.w3c.es/divulgacion/guiasbreves/HojasEstilo>

2.- Representan un lenguaje artificial usado para definir la presentación de un documento estructurado escrito en HTML o XML

Curso avanzado para la construcción de Bibliotecas Digitales utilizando el software Greenstone : aspectos técnicos y desarrollo. Prodigio Consultores Noviembre 2009 . Cuba

HTML:

1. Es el lenguaje con el que se definen las páginas Web. Básicamente se trata de un conjunto de etiquetas que sirven para definir el texto y otros elementos que compondrán una página Web. El HTML es un lenguaje de marcación de elementos para la creación de documentos hipertexto. El archivo donde está contenido el código HTML es un archivo de texto, con una peculiaridad, que tiene extensión .html o .htm (es indiferente cuál utilizar).

   Álvarez M.A. Llegó el momento de hablar sobre HTML, el lenguaje con el que crean las

páginas Web. [Sitio Web en Internet]. 2001/09/29. Disponible en: <http://www.desarrolloweb.com>

2. El HTML, acrónimo inglés de HyperText Markup Language (lenguaje de marcas hipertextuales), lenguaje de marcación diseñado para estructurar textos y presentarlos en forma de hipertexto, que es el formato estándar de las oaginas Web. Gracias a Intenet y a los navegadores Web del tipo Internet Explorer, Opera, Firefox o Netscape, el HTML se ha convertido en uno de los formatos más populares que existen para la construcción de documentos y también de los más fáciles de aprender. HTML es una aplicación de SGML conforme al estándar internacional ISO 8879. XHTML es una reformulación de HTML 4 como aplicación XML 1.0, y que supone la base para la evolución estable de este lenguaje. XHTML permite la compatibilidad con los agentes de usuario que ya admitían HTML 4 siguiendo un conjunto de reglas.

   PergaminoVirtual. Glosario [Sitio Web en Internet] 2007-06-12 ,Disponible en: <http://www.pergaminovirtual.com.ar/definicion/default.css>

ImagemagicK:

Es una robusta colección de herramientas y bibliotecas Open Source cuya licencia permite consultar, escribir y manipular una o varias imagenes (en 88 grandes formatos) incluyendo los formatos TIFF, JPEG, PNG, PDF, Photo CD, y GIF. ImageMagick es una aplicacion portable y compila bajo casi todos los sistemas que corren sobre 32-bit o 64-bit CPUs. ImageMagick se encuentra disponible para Unix o sistemas semejantes incluyendo Linux. Tambien corre sobre Winsows 95,98, ME, NT 4.0, 2000, and XP), Macintosh (9 y 10), VMS, and OS/2. http://www.imagemagick.org

Imagemagick.org. imagemagick. [Website in Internet] 2006. Available from:<http://www.imagemagick.org >

Interoperabilidad:

Desde un punto de vista informático, interoperabilidad se define como la habilidad que tiene un sistema o producto para trabajar con otros sistemas o productos sin un esfuerzo especial por parte del cliente. Este concepto tiene una importancia creciente a tenor de las colecciones digitales distribuidas que utilizan distintos esquemas de metadatos. A pesar de la complejidad de este concepto y de sus múltiples implicaciones para los sistemas de recuperación de información basados en metadatos, es un concepto clave al hablar de esquemas de metadatos y de la necesidad de compatibilizar todos ellos, para una recuperación de información integral en distintas colecciones de datos y metadatos distribuidos. La interoperabilidad entre distintos esquemas de metadatos puede realizarse de diversas formas, por ejemplo a través de el funcionamiento de un protocolo tipo OAI o bien a través del mapeo o establecimiento de correspondencias entre informaciones en diferentes formatos (por ej. MARC-DC, FGDC-DC, etc.) para la conversión de elementos de metainformación que permita hacerlos compatibles.

Mendez E. Senso A. Introduccion a los metadatos. Estándares y aplicación Conceptos asociados al uso de meta información . Unidad de Autoformación. [Sitio Web en Internet] SEDIC 2004. Disponible en: <http://www.sedic.es >

Java Runtime Environment:

Para poder usar la Interfaz de Bibliotecario de Greenstone es necesario contar con una versión adecuada del Ambiente de Ejecución Java (JRE por sus siglas en inglés).

Greenstone.org. Guía de Instalación Greenstone. Apendice: Programas Asociados.[Sitio Web

en Internet]. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Install_es_all.html#Chapter_versions_of_greenstone >

Lenguaje de Macros:

1. El lenguaje de macros, se utiliza para conferir al sistema de Biblioteca Digital Greenstone un estilo coherente y crear interfaces en diferentes lenguas. La interacción con la biblioteca genera la estructura básica de las páginas Web y las macros del directorio GSDLHOME/macros aportan el contenido.

   Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3. El Sistema de Ejecución de Greenstone [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/style.css>

2. Las Macros definen tanto las cadenas de formato de las interfases como las variables utilizadas en Codigo C++.Algunas macros tienen funciones muy singulares como home.dm, query.dm, document.dm o sea que affectan a un solo grupo de elementos mientras otras tienen funcionas globales por ejemplo base.dm y las relacionadas con los idiomas english.dm, french.dm y otras. A una macro se le reconoce por su estructura, su nombre empieza y termina con un guion bajo y encierra su contenido entre llaves como se muestra en el siguiente ejemplo:_nombre de macro_{contenido}

   Greenstone .org. Biblioteca Digital Greenstone. Guía del Programador .Configuración del sitio Greenstone. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/es/html/Chapter_configuring_your_site.htm >

Lenguaje PHP:

PHP es un lenguaje de programación muy potente que, junto con html, permite crear sitios Web dinámicos. Php se instala en el servidor y funciona con versiones de Apache, Microsoft IIs, Netscape Enterprise Server y otros. La forma de usar php es insertando código php dentro del código html de un sitio Web. Cuando un cliente (cualquier persona en la Web) visita la página Web que contiene éste código, el servidor lo ejecuta y el cliente sólo recibe el resultado. Su ejecución, es por tanto en el servidor, a diferencia de otros lenguajes de programación que se ejecutan en el navegador.

masadelante.com. ¿Que significa php? - Definición y explicación del término php. [Sitio Web en Internet] 1999. Disponible en:<http://www.masadelante.com/favicon.ico >

Metadato:

1. El término fue acuñado por Jack Myers en la década de los 60 para describir conjuntos de datos. La primera acepción que se le dió (y actualmente la más extendida) fue la de dato sobre el dato, ya que proporcionaban la información mínima necesaria para identificar un recurso.

   Senso J. A. de la Rosa A. El concepto de metadato. Algo más que descripción de recursos electrónicos. [Sitio Web en Internet]. 2003<http://www.scielo.br/pdf/ci/v32n2/17038.pdf >

2. Los metadatos, es decir, las informaciones descriptivas tales como el autor, el título, la fecha, las palabras clave, etc., pueden asociarse a cada documento o a cada sección de documento. Es importante destacar que los metadatos constituyen la materia prima de los índices de consulta, es decir los formatos que ordena Greenstone para posibilitar la recuperación de documentos. Se los debe suministrar explícitamente o deben poder

extraerse automáticamente de los documentos de origen.

Greenstone Wiki. Biblioteca Digital Greenstone. Guía de Usuario. Características del Programa [Sitio Web en Internet], Marzo 2006. Disponible en: <http://greenstonewiki.cs.waikato.ac.nz/wiki/gsdoc/GNUFDL.html >

3. Son ejemplos de metadatos:

   El encabezamiento de un fichero multimedia (imagen, vídeo o audio).
   El resumen de un documento.
   El catálogo de una base de datos.
   Los términos asignados haciendo uso de un tesauro.
   Las palabras extraídas de un texto.
   Las fichas catalográficas en cualquier formato (ISBD, MARC, etc.).
   Las páginas amarillas.

Mendez E. Senso A.Introduccion a los metadatos. Estándares y aplicación Conceptos asociados al uso de meta información . Unidad de Autoformación.[Sitio Web en Internet] SEDIC 2004. Disponible en: <http://www.sedic.es >

METS (Metadata Encoding and Transmission Standard):

Estándar para codificar una descripción, administrativa, y estructural de los metadatos, con relación a los objetos de una biblioteca digital, expresados con el lenguaje XML y el Consorcio 3W. El estándar es mantenido por Network Development and MARK Standards Office de la Biblioteca del Congreso ,y ha sido desarrollado como una iniciativa de la Digital Library Federation.

Library of Congress. Metadata Encoding and Transmission Standard (METS) Official Web Site [Website in Internet], November 25, 2009. Available from: <http://www.loc.gov/standards/mets/mets-home.html >

Metadatos AGLS (Australian Government Locator Service):

1.- AGLS (Australian Government Locator Service) y NZGLS (New Zealand Government Locator Service) son esquemas de metadatos específicos para la descripción y recuperación de registros documentales procedentes de las Administraciones o de las relaciones burocráticas electrónicas en Australia y Nueva Zelanda, respectivamente. En realidad, AGLS y NZGLS son modelos de metadatos basados en el Dublin Core, al que se le han añadido elementos específicos aplicables a la información gubernamental.

Méndez E. La descripción de documentos electrónicos a través de metadatos: una visión para la Archivística desde la nueva e-Administración [Seriada en Internet] Revista d’Arxius, 2003, pp. 47-82.

2-El estándar de metadatos AGLS es un juego de propiedades descriptivas que el gobierno, departamentos y agencias pueden usar para posibilitar la visibilidad y accesibilidad de sus servicios Web y enlazar los datos de las aplicaciones . Ha sido orientado para el uso de las Agencias estatales australianas. Y se reconocen como AS 5044:2010.

National Archives of Australia. AGLS - National Archives of Australia. AGLS Metadate Standard [Website in Internet] 30/03/2007. Modified 20/08/2010. Available from:<http://www.naa.gov.au/records-management/create-capture-describe/describe/agls/index.aspx >

Metadatos Dublin Core:

La norma ISO15836 define el Conjunto de Elementos Dublin Core, o lo que se conoce habitualmente como "DC simple". Esos 15 elementos básicos para describir cualquier objeto de información, se presentan habitualmente divididos en tres grupos que indican la clase o alcance de la información incluida en ellos, y que responden, en cierta medida, a las expectativas que tiene el usuario cuando se enfrenta a la información de la Red:

Contenido	Propiedad intelectual	"Instantiation"/ Ejemplo
Title	Creator	Date
Subject	Publisher	Type
Description	Contributor	Format
Source	Rights	Identifier
Language		
Relation		
Coverage		

Mendez E. Senso A.Introduccion a los metadatos. Estándares y aplicación. Uso del Dublin Core. Unidad de Autoformacion . .[Sitio Web en Internet] SEDIC 2004. Disponible en: <http:www.sedic.es>

Metadatos RFC 1807:

Formato para los registros bibliográficos desarrollado por R. Lasher & D. Cohen, (Junio 1995). Es un memo, no un verdadero estándar de metadato y define un formato para los registros bibliográficos en el correo electrónico de los reportes técnicos. Es usado por la Universidad de Cornell, la que cuenta con una biblioteca digital abierta que utiliza el protocolo Dienst y por la universidad de Stanford con el sistema SIFT (para nuevos grupos).

RFC 1807 también es utilizado por las cinco instituciones para la ciencia de la computación del proyecto ARPA (Cornell, stanford. UC, MIT y la Universidad de Carnegie Mellon

UKOLN Metadata Group. *A review of metadata: a survey of current resource description formats* – RFC1807 [Website in Internet] Work Package 3 of Telematics for Research project DESIRE (RE 1004) 10-Jun-19998. Disponible en: <http://www.ukoln.ac.uk/metadata/ >

MG (abreviación de Managing Gigabytes):

1. Greenstone utiliza MG (abreviación de Managing Gigabytes) para indizar y recuperar documentos, y su código fuente se incluye en el directorio GSDLHOME/packages. MG emplea técnicas de compresión para optimizar la utilización del espacio de disco sin alterar la rapidez de ejecución. En el caso de una colección de documentos en inglés, el texto comprimido y los índices suelen ocupar un tercio del espacio que necesitaría el texto original sin comprimir. La búsqueda y recuperación suelen ser más rápidas que las operaciones equivalentes en la versión sin comprimir porque requieren menos operaciones en el disco duro.

   Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. . Capitulo 3.El Sistema de Ejecución de Greenstone [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css >

2. Programa utilizado en el sistema Greenstone para la indexación en texto completo, que incorpora técnicas de compresión.

   Witten, I.H; Mofft, A. y Bell, T Managing Gigabytes: compressing and indexing documents

and images, Morgan Kaufmann, segunda edición, 1999).

Necesidades Informativas:

1. Para establecer las necesidades informativas es necesario considerar el carácter objetivo que presentan al existir independientemente de la conciencia de quien la tiene.

   Rodríguez I. Rodríguez R. Estudio de las necesidades informativas de los profesionales e investigadores del Sistema Nacional de Salud. ACIMED 21 (1):47-54.

2. Es necesario conjugar la imagen subjetiva que tiene el usuario de la información, es decir sus intereses de información y comparar por demás ésta con sus necesidades objetivas; determinando si hay coincidencia o no para de ese modo entonces hacerlo consciente de las mismas.

   Núñez I. Guía metodológica para el estudio de las necesidades de formación e información de los usuarios o lectores. ACIMED 5 (3)32-5 sept-dic: Disponible en <http://bus.sld.cu/revistas/aci/vol05_03_97 >

OAI-PMH (Protocole Open Archives -Protocol para Metadata Harvesting):

OAI-PMH es un protocolo de recopilación con un enfoque de interoperabilidad que proporciona una búsqueda distribuida que equivale a buscar y descubrir información y servicios remotos y recopilación, lo que equivale a que los datos/metadatos son transferidos desde la fuente remota al destino en el cual realizaran los servicio de búsqueda. OAI-PMH funciona sobre el protocolo HTTP. Las peticiones son operaciones HTTP GET o POST. Las respuestas son documentos XML válidos.

OAI-Iniciativa que pretende:

1. desarrollar y promover estándares de interoperabilidad para la difusión de contenidos en Internet
2. mejorar el acceso a publicaciones electrónicas (eprints) y
3. abrir el acceso a un rango de materiales digitales (objetivo más amplio).

Van de Sompel, H., Lagoze C. The Santa Fe convention of the Open Archives Initiative D-Lib [Serials in Internet] . 2000, Febrero, v. 6, n. 2.Available from:<http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html>

OpenDOAR:

Es un Directorio Académico de Repositorios Open Acces. Cada uno de estos repositorios han sido revisados para observar la calidad de los mismos. Cuenta además con herramientas que soportan la gestión administrativa y las mejores practicas de uso en su infraestructura.

OpenDOAR - Home Page - Directory of Open Access Repositories. [Website in Internet] 11-Sep-2008 http://www.nottingham.ac.uk/

Perl (Practical Extraction and Report Language):

1. Lenguaje de programación utilizado en el WWW a través de un CGI, principalmente para realizar consultas a bases de datos como Oracle, SQL-Server, SyBase, etc, o a herramientas locales como WAIS. Perl es un lenguaje para manipular textos, archivos y procesos, proporciona una forma fácil y legible para realizar trabajos que normalmente se realizarían en Co en un shell. Perl nació y se ha difundido bajo el sistema operativo UNIX, aunque existe para otras plataformas.

Profesional Hosting. PERL [Website in Internet] 2002. Disponible en:<http://www.profesionalhosting.com/hosting/definicion/style.css >

2. Es un lenguaje de programación creado por Larry Wall en 1987. Perl toma características del C, del lenguaje interpretado shell (sh), AWK, sed, Lisp y, en un grado inferior, muchos otros lenguajes de programación. Estructuralmente, Perl está basado en un estilo de bloques como los del C o AWK, y fue ampliamente adoptado por su destreza en el procesado de texto y por no tener ninguna de las limitaciones de los otros lenguajes de script.

   PergaminoVirtual.com. Definición de PERL [Sitio Web en Internet] 2007-06-12 <http://www.pergaminovirtual.com.ar/definicion/default.css>

Phind:

Programa autónomo incluido en Greestone que representa una herramienta de consulta jerárquica en los grupos de palabras.

Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. . Capitulo 3. El sistema de Ejecución de Greenstone [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css>

Ping (Packet Internet Groper-Rastreador de Paquetes Internet):

1. Se refiere a la utilidad que comprueba el estado de la conexión con uno o varios equipos remotos por medio del envío de paquetes de solicitud de eco. Entre otras aplicaciones, el ping resulta útil para medir el tiempo de conexión entre dos puntos remotos.

   PergaminoVirtual.com. Definición de PING. [Sitio Web en Internet. 11/08/10 Disponible en: <http://www.pergaminovirtual.com.ar/definicion/default.css >

2. Ping (Packet Internet Groper) Proceso de verificación de la oportunidad de una computadora o del buen funcionamiento de una conexión de red. Se realiza introduciendo el comando **ping 127.0.0.1** para verificar la propia computadora pc, o el IP del ordenador cuya comprobación deseamos hacer.

   Marketmagazine. Definición de Ping [sitio Web] (consultado 6 de febrero de 2005). Disponible en <http://www.marketmagazine.bosfo>

Pluggins:

1. Un plug-in es un módulo de hardware o software que añade una característica o un servicio específico a un sistema más grande.La idea es que el nuevo componente se enchufa simplemente al sistema existente. Por ejemplo, hay un gran número de plug-ins para el navegador Firefox que te permiten utilizar diversas herramientas y el plug-in de Flash, permite ver animaciones en flash en cualquier navegador.

   Mas Adelante.com. Qué es un pluging?. Definición de Plugging [sitio Web] 1999. Disponible en: http://www.masadelante.com/faqs/plug-in

2. Plug-in es un programa que extiende las capacidades del navegador de un modo especifico, dado por ejemplo la capacidad de mostrar vídeo, audio, ficheros de un determinado formato (ficheros PDF, presentaciones de ASAP, fichero VRML, etc …). No existe actualmente un conjunto estándar de plug-in´s para cada tipo de ficheros, sino que existen diversas aplicaciones, realizadas por diversos fabricantes, y no todas de libre distribución. Se puede asegurar que todas las aplicaciones serán compatibles y si por ejemplo se referencia un fichero de sonido en formato .wav en su página, este podrá ser oído por todos aquellos que

tengan un plug-in para este tipo de ficheros. Por tanto se recomienda un uso moderado de los plug-in´s utilizando ficheros que sean de formatos comunes y un uso complementario en una página, ya que muchos usuarios no podrán verlo. Es aconsejable incluir un enlace en el que se referencie el fichero, de esta forma un usuario podrá seguir el enlace y ver el fichero en un visualizador externo si lo posee.

Masternagazine. Definición de Plug-in.[Sitio Web en Internet] 16 de Febrero de 2005. Disponible en: <http://www.mastermagazine.info>

3. Los conectores analizan los documentos importados y extraen los metadatos en Greenstone . El conector HTML, por ejemplo, convierte las páginas HTML en el Formato de Archivo Greenstone y extrae los metadatos que están explícitos en el formato del documento, como los títulos entre etiquetas <title> </title>. Los conectores se escriben en lenguaje Perl. Todos proceden de un conector básico denominado BasPlug, que realiza todas las operaciones necesarias como crear un nuevo documento en el Formato de Archivo Greenstone con que trabajar, asignar un identificador de objeto (OID) y manejar las secciones de un documento. Los conectores se guardan en el directorio perllib/plugins.

   Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 2. Explotación Óptima de sus Documentos. [Sitio Web en Internet] 2004. Disponible en: <http://www.greenstone.org/manuals/gsdl2/style.css >

Servidor:

En informática, un servidor es un tipo de software que realiza ciertas tareas en nombre de los usuarios. El término servidor ahora también se utiliza para referirse al ordenador físico en el cual funciona ese software, una máquina cuyo propósito es proveer datos de modo que otras máquinas puedan utilizar esos datos. Este uso dual puede llevar a confusión. Por ejemplo, en el caso de un servidor Web, este término podría referirse a la máquina que almacena y maneja los sitios Web, y en este sentido es utilizada por las compañías que ofrecen hosting o hospedaje. Alternativamente, el servidor Web podría referirse al software, como el servidor de http de Apache, que funciona en la máquina y maneja la entrega de los componentes de las paginas Web como respuesta a peticiones de los navegadores de los clientes.

Pergamino Virtual.com. Definición de Servidor [Sitio Web en Internet]. 11/08/10. Disponible en: <http://www.pergaminovirtual.com/definicion/default.css >

STL (Biblioteca Standard de Plantillas):

Greenstone utiliza considerablemente la Biblioteca Estándar de Plantillas (STL), una biblioteca C++ de amplia difusión concebida por Silicon Graphics (http://www.sgi.com), y que es el resultado de numerosos años de investigación y desarrollo. Como todas las bibliotecas de programación requiere un tiempo de aprendizaje. Para una descripción más completa, consulte el manual oficial de referencia de STL disponible en la dirección Internet http://www.sgi.com

Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3.El Sistema de Ejecución de Greenstone. [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css >

SGML:

El SGML (Standard Generalized Markup Language) es la especificación "padre" de todos los lenguajes de marcado. Es un estándar internacional (de la ISO) para la definición de métodos independientes de dispositivos y plataformas de representación de textos en forma electrónica.

Es también una estructura que permite guardar información que facilita la manipulación, búsqueda precisa y el rearreglo de grandes repositorios de información. El SGML es una especificación general y presenta el concepto de tipo de documento. Así, cada tipo de documento tiene una especificación de estructura o DTD (Documento Type Definition). El SGML puede especificar "desde transcripciones de antiguos rollos sumerios hasta la documentación técnica de bombarderos Stealth y desde historias clínicas de pacientes hasta notas musicales", y por supuesto, incluye al hipertexto.

Gaiasur.com. Los lenguajes de marcado [Sitio Web en Internet], 2006. Disponible en: <http://fractal.gaiasur.com.ar/infoteca/seminarios/pub-/conceptos/hipertexto.html >

Sociedad Informacional:

1. La sociedad de la información es un estadío de desarrollo social caracterizado por la capacidad de sus miembros (ciudadanos, empresas yadministración pública) para obtener y compartir cualquier información,instantáneamente desde cualquier lugar y en la forma que se prefiera.

   Bangemann M. Europa i la societat global de la información. Recomanacions al Consell Europeu. [Website in Internet] Brussel.les, 26 de maig de 1994. Available from: <www.sre.urv.es/Web/pled/.../pla/.../01_informe_bangemann.pdf >

2. Nuevo sistema tecnológico,económico y social. Una economía en la que el incremento de productividad no depende del incremento cuantitativo de los factores de producción (capital,trabajo,recursos naturales),sino de la aplicación de conocimientos e información a la gestión,producción y distribución tanto en los procesos como en los productos.

   Castells M. The Information Age: Economy, Society and Culture. [Website in Internet] 1998. Available from: <http://sociology.berkeley.edu/faculty/castells/trilogy_reviews.html>

3. Forma específica de organización social en la que la generación, proceso, y transmisión de información, se convierten en las principales fuentes de productividad y poder.

   Cornella, A. ¿Economía de la información o sociedad de la información?. Barcelona. ESADE, Marzo, 1994.

Sistema CORBA:

1. CORBA define la infraestructura para la arquitectura OMA (Object Management Arquitecture) de OMG, especificando los estándares necesarios para la invocación de métodos sobre objetos en entornos heterogéneos. Los entornos heterogéneos son aquellos en los que las arquitecturas que constituyen el entorno pueden ser sistemas Microsoft Windows, máquinas Unix de diferentes fabricantes (Linux entre otros) e inclusos sistemas como MacOS o OS/2. Y es más, dentro de la heterogeneidad también se incluyen los sistemas de comunicaciones (protocolos de comunicación como TCP/IP, IPX ...) o los lenguajes de programación utilizados en las diferentes arquitecturas. Es el mundo informático en su complejidad más alta y por tanto sólo una arquitectura muy flexible y potente puede cubrir todos estos aspectos. CORBA define su propio modelo de objetos, basado en la definición de las interfaces de los objetos mediante el lenguaje IDL. CORBA ha buscando un entorno heterogéneo, el cual constituye una visión abierta del mundo de la informática y en la cual hay cabida para diferentes sistemas y distintas filosofías, un mundo más rico que el que se puede lograr con un solo sistema alrededor del cual funcionan todas las aplicaciones.

   Programación en Castellano,S.L..Curso práctico de Corba en GNU/Linux. [Sitio Web en

Internet].2008.Disponible en: <http://www.programacion.com/templates/default/css/estilo.css >

2. CORBA (Common Object Request Broker Arquitecture) CORBA al igual que varias tecnologías aceptadas hoy en día es solo una especificación que fue creada en 1989 por OMG (Object Management Group). Como el nombre de la organización lo implica, CORBA establece estándares para la comunicación de objetos a través de procedimientos/métodos remotos. IDL ("Interface Definition Language"). IDL es un lenguaje utilizado para crear cualquier desarrollo en CORBA, su nombre es un indicador de su funcionamiento: definición de interfases, esto es, a través de IDL se definen las diversas estructuras que serán utilizadas en un ambiente CORBA.

   Osmosislatina.com. RMI y CORBA: Una base para aplicaciones Distribuidas. [Sitio Web en Internet] 2005/09/07. Disponible en: <http://www.osmosislatina.com/css/osmosis2006.css >

3. Utiliza un paradigma unificado orientado a objetos, para permitir a distintos procesos ejecutárse en distintas plataformas informáticas e implementarse en diferentes lenguajes de programación, acceder a la misma serie de objetos distribuidos por Internet (o cualquier otra red). Con el sistema CORBA, se puede acceder a un servidor remoto de colección de Greenstone, escrito en C++.

   Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3.El Sistema de Ejecución de Greenstone. [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css >

Software Libre:

Se refiere a la libertad de los usuarios para ejecutar, copiar, distribuir, estudiar, cambiar y mejorar el software. Esto se concreta en cuatro atributos:

- La libertad de usar el programa, con cualquier propósito (libertad 0).
- La libertad de estudiar cómo funciona el programa, y adaptarlo a tus necesidades (libertad 1). El acceso al código fuente es una condición previa para esto.
- La libertad de distribuir copias, con lo que puedes ayudar a tu vecino (libertad 2).
- La libertad de mejorar el programa y hacer públicas las mejoras a los demás, de modo que toda la comunidad se beneficie. (libertad 3).
- El acceso al código fuente es un requisito previo para esto.

- Free Software Foundation. La Definición de Software Libre. [Website in Internet]. Proyecto GNU. Free Software Foundation. 2006/05/20. Disponible en: <http://www.gnu.org/philosophy/free-sw.es.html>

Software de código fuente abierto:

Describe una categoría de licencias de software casi, pero no completamente, igual que "software libre". El Software de código fuente abierto incluye software libre, pero también incluye programas semi-libres, e incluso algunos programas privativos bajo su licencia inicial.

Stallman R. Porqué "Software Libre" es mejor que software de "Código Fuente Abierto". [Website in Internet]. Free Software Foundation. 6 de noviembre de 1999. Disponible en: < http://www.gnu.org/philosophy/free-software-for-freedom.html>

Software Apache:

Servidor open source para los modernos sistemas operativos que incluye UNIX y Windows NT. El objetivo es proveer un servidor http , seguro, eficiente, y extensible con los estándares

La Habana : Centro de Cibernética Aplicada a la Medicina (CECAM) - Editorial Universitaria, 2013. -- ISBN 978-959-16-1494-0

http vigentes.

Apache Http server project. The number one HTTP server on the Internet. [Website in Internet] 2010/04/27. Disponible en: <http://httpd.apache.org> and <http://www.apache.org>

Tercer Capitalismo o Capitalismo Cognitivo:

El cual se caracteriza por colocar en primer plano la redefinición de los derechos de propiedad, en una sociedad digital y manejar también el concepto de activos intangibles como motor de desarrollo.

Boutang Y. Riqueza, propiedad, libertad y renta en el capitalismo cognitivo en: Capitalismo cognitivo, propiedad intelectual y creación colectiva. [Libro en Internet]Edición: Traficantes de Sueños Madrid; Abril de 2004. Disponible en: <http://traficantes.net7>.

Txt2db/:

Convierte un formato de texto ASCII de tipo XML en formato de base de datos de GNU.

Grenstone.org. Biblioteca Digital Greenstone; Guía del Programador. Capitulo 3.El Sistema de Ejecución de Greenstone. [Sitio Web en Internet] 2004. Disponible en:<http://www.greenstone.org/manuals/gsdl2/style.css >

Unicode:

1. Sistema que permite procesar cualquier idioma de forma sistemática. Antes de unicode, había diferentes sistemas de codificación de caracteres para cada idioma, cada uno usando los mismos números (0-255) para representar los caracteres de ese lenguaje. Intercambiar documentos entre estos sistemas era difícil porque no había manera de que un computador supiera con certeza qué esquema de codificación de caracteres había usado el autor del documento; el computador sólo veía números, y los números pueden significar muchas cosas. Para resolver estos problemas, unicode representa cada carácter como un número de 2 bytes, de 0 a 65535. Cada número de 2 bytes representa un único carácter utilizado en al menos un idioma del mundo (los caracteres que se usan en más de un idioma tienen el mismo código numérico). Hay exactamente 1 número por carácter, y exactamente 1 carácter por número. Los datos de unicode nunca son ambiguos.

   Pilgrim M. Inmersión en Python. Capítulo 9. Procesamiento de XML. 9.4. Unicode.2005. [Libro en Internet]. Disponible en: <http://www.gulic.org/almacen/diveintopython-5.4-es/diveintopython.css>

2. La Biblioteca del Congreso de Estados Unidos ha designado la norma ISO 639-2/RA con el objetivo de procesar las consultas para el lenguaje alpha-3 integrado en los Estándares Internacionales para la representación de los nombres de los idiomas.

   Library of Congress. Part 2: alpha-3 code. ISO 639-2 Registration Authority - [Website in Interne].October 20, 2009. Available from: <http://www.loc.gov/standards/iso639-2/iso639-2ra.html >

Usabilidad:

La ISO define la usabilidad como la capacidad que tiene un producto para ser usado por determinados usuarios con el fin de alcanzar unos objetivos concretos con efectividad, eficiencia y satisfacción dentro de un contexto de uso específico (ISO 9241-11). Esta definición resulta muy clarificadora, pues divide la usabilidad en los diversos aspectos que considera: efectividad, eficiencia y satisfacción; y hace una diferenciación en función del

objetivo, el usuario y el contexto de uso. La usabilidad no es un valor absoluto e intrínseco a un sistema, sino que puede variar en función del contexto. La definición de la ISO también deja claro que un diseño se realiza para unos usuarios y objetivos determinados. Por lo tanto, no puede afirmarse que un sistema sea usable o no en sí mismo, sino que habrá que concretar las variables mencionadas.

Usabilidadweb. Definiciones ISO Modelo de Proceso de la Ingeniería de la usabilidad y de la accesibilidad. 2005/10/26 MPIu+a GRIHO, 2005. Disponible en: <http://usabilidadweb.htm>

Atributos de Usabilidad: facilidad de aprendizaje, eficiencia, recuerdo en el tiempo, tasa de errores y satisfacción. A continuación se explican detalladamente:

a) **Facilidad de aprendizaje:** Cuán fácil es aprender la funcionalidad básica del sistema, como para ser capaz de realizar correctamente la tarea que desea realizar el usuario. Se mide normalmente por el tiempo empleado con el sistema hasta ser capaz de realizar ciertas tareas en menos de un tiempo dado (el tiempo empleado habitualmente por los usuarios expertos). Este atributo es muy importante para usuarios noveles.

b) **Eficiencia**: El número de transacciones por unidad de tiempo que el usuario puede realizar usando el sistema. Lo que se busca es la máxima velocidad de realización de tareas del usuario. Cuanto mayor es la usabilidad de un sistema, más rápido es el usuario al utilizarlo, y el trabajo se realiza con mayor rapidez. Nótese que eficiencia del software en cuanto su velocidad de proceso no implica necesariamente eficiencia. del usuario en el sentido en el que aquí se ha descrito.

c) **Recuerdo en el tiempo**: Para usuarios intermitentes (que no utilizan el sistema regularmente) es vital ser capaces de usar el sistema sin tener que aprender cómo funciona partiendo de cero cada vez. Este atributo refleja el recuerdo acerca de cómo funciona el sistema que mantiene el usuario, cuando vuelve a utilizarlo tras un periodo de no utilización.

d) **Tasa de errores**: Este atributo contribuye de forma negativa a la usabilidad de un sistema. Se refiere al número de errores cometidos por el usuario mientras realiza una determinada tarea. Un buen nivel de usabilidad implica una tasa de errores baja. Los errores reducen la eficiencia y satisfacción del usuario, y pueden verse como un fracaso en la transmisión al usuario del modo de hacer las cosas con el sistema.

e) **Satisfacción:** Éste es el atributo más subjetivo. Muestra la impresión subjetiva que el usuario obtiene del sistema. Algunos de estos atributos no contribuyen a la usabilidad.

Nielsen, J. Usability Engineering. AP Professional, 1993. Citado por: Mari-Carmen Marcos et al.. Evaluación de la usabilidad en sistemas de información terminológicos online [Seriada en Internet]. "Hipertext.net", núm. 4, 2006. Disponible en: <http://www.hipertext.net>

XML:

1. El lenguaje extensible de marcas, abreviado XML, describe una clase de objetos de datos llamados documentos XML y parcialmente describe el comportamiento de programas de computador que pueden procesarlos. XML es un perfil de aplicación o forma restringida de SGML (Standard Generalized Markup Language) [ISO 8879]. Por construcción, todo documento conforme con XML es conforme con SGML. Los documentos XML están hechos de unidades de almacenamiento llamadas entidades, las cuales contienen datos procesados (parsed) o sin procesar. Los datos procesados están hechos de caracteres, algunos de los cuales forman datos de carácter y otros marcas. Las marcas codifican la

descripción del esquema de almacenamiento y estructura lógica del documento. XML provee un mecanismo para imponer restricciones al esquema de almacenamiento y estructura lógica. El esquema XML es un dialecto simple de SGML diseñado para hacer posible que el genérico SGML sirva, reciba y procese en el Web en el mismo sentido que ahora es posible con el HTML . El XML ha sido diseñado para facilitar la implementación interoprabilidad con SGML y HTML.

W3C El lenguaje extensible de marcas (XML) 1.0 - Extensible Markup Language (XML) Recomendación de la W3C [Website in Internet], - Febrero 1998. Editores: Tim Bray (Textuality y Netscape) Jean Paoli (Microsoft) C. M. Sperberg- McQueen (University of Illinois at Chicago) . Traducido al español por: Fabio Arciniegas A. 1999/08/17. Disponible en: <http://www.sidar.org >

2. El esquema XML comparte el vocabulario y conduce a las maquinas a seguir determinadas reglas. Provee un medio para definir la estructura , contenido y semántica de los documentos XML. Fue aprobado por el consorcio 3W en una recomendación del 2 de mayo de 2001.

   W3C. XML Schema Working Group XML Schema User Experience Workshop, [Website in Internet], 21-22 June 2005 Available from: <http://www.w3.org/XML/2009/02/schema-charter.html >

Z39.50:

1. Z39.50 es una norma para consultar catálogos de bibliotecas en Internet utilizando las mismas reglas para todos.los catálogos, algo así como usar siempre la misma pantalla para consultar cualquier biblioteca. Para consultar sistemas Z39.50 es necesario contar con un programa especial llamado cliente Z39.50 o navegador Z39.50, o bien a través de páginas o de portales en Internet. Es recomendable contar con un programa de consulta porque no todas las bibliotecas son accesibles a través de WWW.

   Biblioteca Daniel Cosío Villegas. Bibliotecas en Z39.50 [Sitio Web en Internet]. 2000. Disponible en: <http://biblio.colmex.mx/portal/estela14.JPG >

2. Z39.50 es un estándar internacional para la comunicación entre computadoras , bibliotecas y sistemas de información relacionados. Fue propuesto originalmente en 1984 con el objetivo de proporcionar una vía estándar de interrogación a las bases de datos bibliográficas . es mantenido por la Agencia Z39.50, que es administrada por la Biblioteca del congreso de Estados Unidos.

   Library of Congress. Biblio Tech Review. Z39.50: Part 1- an overview 12 [Website in Internet]. April, 2001. Available from: <http://lcweb.loc.gov/z3950/ >

ANEXOS

Anexo Construyendo una colección de multimedia (instrucciones básicas)

1 Comenzar una nueva colección en base a:

2 Copiar los files de → *advbeat_small*

3 Aparecerán unos mensajes automáticamente que les preguntaran si quieren añadir *pluggins* específicos para MP3 y MARC, pues estos formatos existen en los documentos que hemos copiado.

4 Cambie al panel Enriquecer y asigne valores a los metadatos seleccionados.

5 Dirijase al panel Crear y comience a elaborar la colección

6 Puede revisar con el botón vista previa y vera la colección desde la interfaz de usuario con todas sus características. También puede revisar la existencia de la colección creada si se dirige a al botón file y hace clic sobre open. Aparecerá la relación de las colecciones disponibles desde el GLI.

7 Si quisieran modificar los valores asignados a los clasificadores seleccionados pueden hacerlo manualmente, dirigiéndose al panel Diseño y el botón *Browsing Classiffiers* y haciendo doble clic sobre el dato para que aparezca la ventana de dialogo. O también un clic en la ventana de abajo denominada *Configure Clasiffiers.*

8 El próximo ejercicio será incorporar el clasificador PHIND. Para ello desde el panel Diseño y *Browsing Clasiffiers* agregar el clasificador PHIND. Cree nuevamente la colección y en la vista previa vera como queda añadido el clasificador.

9 Insertemos ahora una imagen que caracterice la colección. Ir al panel de formato en la sección General y allí utilizar el botón URL to *about page' image:* y mediante la opción de *browse* que posee buscar la imagen que queremos utilizar para insertar. Cree de nuevo la colección y vera como aparece la imagen

10 Introducción en la colección el UnknownPlugin;. Iremos al Panel Diseño y utilizando la opción Document Plugin agregaremos este plugin con el botón *Add Plugin* que encontramos abajo. Cuando estemos configurando los argumentos para este plugin activaremos la casilla *process-extension* y allí escribiremos mid para que este plugin nos ayude a procesar los formatos midi, Activaremos también la casilla File-Format y escribremos allí audio/midi.Cree de nuevo la colección y vera como puede reconocer ahora los archivos midi.

Greenstone es un sistema de código abierto (Open Source) creado por la Universidad de Waikato, con el propósito de facilitar a los interesados, la creación de bibliotecas digitales o repositorios institucionales.

Los autores de este libro han intentado volcar en el, toda la experiencia acumulada en el desarrollo de su propio proyecto de creación de una biblioteca digital, y la experiencia alcanzada como docentes, en el dictado de cursos básicos y avanzados en esta materia, con el objetivo de promover la utilización del mismo en Cuba y en Latinoamérica como parte del Proyecto Regional "*Grupo de Usuarios Greenstone de Latino-américa*"

La gran actividad demostrada por el Centro de Cibernética Aplicada a la Medicina (CECAM) en pos de la difusión de Greenstone no hace más que manifestar el profundo deseo que tienen por hacer del acceso a la información uno de los pilares de la sociedad.

Estimamos que esta obra en español, puede constituirse en un valioso aporte docente y práctico.

Lic. Diego Spano

Creador del Grupo de Promoción y soporte

de Greenstone para Latino-américa

Administrador de la lista de usuarios en español



ISBN 978-959-16-1494-0